MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

Army Science Conference Proceedings

19-21 June 1984

Volume III

Principal Authors P through Z

This document has been approved for public release and sale; its distribution is unlimited.

DTIC SELECTE DEC 6 1984

D

84 11 30 007

DEPARTMENT OF THE ARMY
OFFICE OF THE DEPUTY CHIEF OF STAFF
FOR RESEARCH, DEVELOPMENT, AND ACQUISITION
WASHINGTON, D.C. 20310

REPLY TO
ATTENTION OF

DAMA-ARR   22 June 1984

SUBJECT: Proceedings of the 1984 Army Science Conference

SEE DISTRIBUTION

1. The fourteenth in a series of Army Science Conferences was held at the
United States Military Academy, 19-21 June 1984. The conference presented a
cross section of the many significant scientific and engineering programs
carried out by the Department of the Army and provided an opportuity for
Department of the Army civilian and military scientists and engineers to
present the results of their research and development efforts before a
distinguished and critical audience.

2. These Proceedings of the 1984 Army Science Conference are a compilation of
all papers presented at the conference and the supplemental papers that were
submitted. The Proceedings consist of five volumes, with Volumes I through
III being unclassified, Volume IV being limited and Volume V being classified.
Papers included in this Volume follow.

3. Our purpose for soliciting these papers was to:

    a. Stimulate the involvement of scientific and engineering talent within
the Department of the Army.

    b. Demonstrate Army competence in research and development.

    c. Provide a forum wherein Army personnel can demonstrate the full scope
and depth of their current projects.

    d. Promote the interchange of ideas among members of the Army scientific
and engineering community.

4. The information contained in these volumes will be of benefit to those who
attended the conference and to others interested in Army research and develop-
ment. It is requested that these Proceedings be placed in technical libraries
where they will be available for reference.

JAMES H. MERRYMAN
Lieutenant General, GS
Deputy Chief of Staff for Research,
    Development and Acquisition

1

Center for Systems Engineering & Integration, ATTN: DRSEL-SEI, Ft. Monmouth, NJ 07703
Center for Tactical Computer Systems, ATTN: DRSEL-TCS, Ft. Monmouth, NJ 07703
US Army Electronics R&D Command, ATTN: DRDEL-CT, Adelphi, MD 20783
US Army Electronics R&D Command, ATTN: DRDEL-SA, Ft. Monmouth, NJ 07703
Atmospheric Sciences Laboratory, ATTN: DELAS-DD, WSMR, NM 88002
Combat Surveillance & Target Acquisition Lab, ATTN: DELCS-D, Ft. Monmouth, NJ 07703
Electronics Technology & Devices Lab, ATTN: DELET-D, Ft. Monmouth, NJ 07703
Electronics Warfare Laboratory, ATTN: DELEW-D, Ft. Monmouth, NJ 07703
Office of Missile Electronic Warfare, ATTN: DELEW-M-D, WSMR, NM 88002
Harry Diamond Laboratories, ATTN: DELHD-TD, Adelphi, MD 20783
Night Vision and Electro-Optics Lab, ATTN: DELNV-D, Ft. Belvoir, VA 22060
Signals Warfare Laboratory, ATTN: DELSW-D, Vint Hill Farms Station, VA 22186
US Army Foreign Science and Technology Center, ATTN: DRXST-OC, Charlottesville, VA 22901
US Army Human Engineering Lab, ATTN: DRXHE-D, Aberdeen Proving Ground, MD 21005
US Army Materiel Systems Analysis Activity, ATTN: DRXSY-D, APG, MD 21005
US Army Materials & Mechanics Rsch Center, ATTN: DRXMR-D, Watertown, MA 02172
US Army Missile Command, ATTN: DRSMI-R, Redstone Arsenal, AL 35898
US Army Research, Development & Standardization Group (Europe), Box 65, FPO New York 09510
US Army Research Office, ATTN: DRXRO-ZC, Rsch Triangle Park, NC 27709
US Army Tank-Automotive Command, ATTN: DRSTA-NS, Warren, MI 48090
Tank-Automotive Systems Lab, ATTN: DRSTA-R, Warren, MI 48090
Tank-Automotive Concepts Lab, ATTN: DRSTA-Z, Warren, MI 48090
US Army Test & Evaluation Command, ATTN: DRSTE-TD, APG, MD 21005
Aberdeen Proving Ground, ATTN: STEAP-PC, APG, MD 21005
Dugway Proving Ground, ATTN: STEDP-SD, Dugway, UT 84022
Electronic Proving Ground, ATTN: STEEP-TD, Ft. Huachuca, AZ 85613
Jefferson Proving Ground, ATTN: STEJP-CO, Madison, IN 47250
Tropic Test Center, ATTN: STETC-CO, APO Miami 34004
White Sands Missile Range, ATTN: STEWS-SC, WSMR, NM 88002
Yuma Proving Ground, ATTN: STEYP-CO, Yuma, AZ 85365
US Army Troop Support Cmd, 4300 Goodfellow Blvd., St. Louis, MO 63120
Belvoir R&D Center, ATTN: STRBE-ZT, Ft. Belvoir, VA 22060
Natick R&D Center, ATTN: STRNC-Z, Natick, MA 01760
US Army Medical R&D Command, ATTN: SGRD-ZA, Ft. Detrick, Frederick, MD 21701
Aeromedical Research Lab, ATTN: SGRD-UAC, Ft. Rucker, AL 36362
Institute of Dental Research, WRAMC, ATTN: SGRD-UDZ, Wash, DC 20012
Inst of Surgical Research, ATTN: SGRD-USZ, Ft. Sam Houston, TX 78234
Letterman Army Inst of Rsch, ATTN: SGRD-ULZ, Presidio of San Francisco, CA 94129
Medical Bioengineering R&D Lab, ATTN: SGRD-UBZ, Frederick, MD 21701
Medical Rsch Inst of Chemical Defense, ATTN: SGRD-UV-ZA, APG, MD 21010
Medical Rsch Inst of Environmental Medicine, ATTN: SGRD-UEZ, Natick, MA 01760
Medical Rsch Inst of Infectious Diseases, ATTN: SGRD-UIZ-A, Frederick, MD 21701
Walter Reed Army Inst of Research, ATTN: SGRD-UWZ, Washington, DC 20307
US Army Health Services Command, Ft. Sam Houston, TX 78234
US Army Research Institute for the Behavioral and Social Sciences, ATTN: PERI-BR, 5001 Eisenhower Avenue, Alexandria, VA 22333
ARI Field Unit, ATTN: PERI-IJ, Ft. Benning, GA 31905
ARI Field Unit, ATTN: PERI-SB, Ft. Bliss, TX 79916

3

ARI Field Unit, ATTN: PERI-SH, Ft. Hood, TX 76544
ARI Field Unit, ATTN: PERI-IK, Ft. Knox, KY 40121
ARI Field Unit, ATTN: PERI-SL, Ft. Leavenworth, KS 66027
ARI Field Unit, ATTN: PERI-IR, Ft. Rucker, AL 36362
ARI Field Unit, ATTN: PERI-SS, Ft. Sill, OK 73503
ARI Field Unit, ATTN: PERI-IO, Presidio of Monterey, CA 93940
US Army Training and Doctrine Command, ATTN: ATDO-ZT, Fort Monroe, VA 23651
  TRADOC Systems Analysis Activity, WSMR, NM 88002
  Training Technology Agency, ATTN: ATTG-D, Ft. Monroe, VA 23651
COMMANDANTS:
  US Army Air Defense School, ATTN: ATSA-CDC-S, Ft. Bliss, TX 79916
  US Army Armor School, ATTN: ATSB-CD, Ft. Knox, KY 40121
  US Army Aviation Logistics School, ATTN: ATSG-DA, Ft. Eustis, VA 23604
  US Army Chemical School, ATTN: ATZN-CM-AS, Ft. McClellan, AL 36205
  US Army Engineer School, ATTN: ATZA-CD, Ft. Belvoir, VA 22060
  US Army Field Artillery School, ATTN: ATSF-FD, Ft. Sill, OK 73503
  US Army Infantry School, ATTN: ATSH-RM, Ft. Benning, GA 31905
  US Army Intelligence Center and School, ATTN: ATSI-SA, Ft. Huachuca, AZ 85613
  US Army Logistics Center, ATTN: ATCL-SCA, Ft. Lee, VA 23801
  US Army Ordnance Center and School, ATTN: ATSL-CD-CS, APG, MD 21005
  US Army Signal Center, ATTN: ATZH-DC/AC, Ft. Gordon, GA 30905
  US Army Soldier Support Center, ATTN: ATZI-DDR, Ft. Harrison, IN 46216-5700
  US Army Transportation School, ATTN: ATSP-AC, Ft. Eustis, VA 23604
PROJECT MANAGERS:
  FIREFINDER/REMBASS, ATTN: DRCPM-FFR, Ft. Monmouth, NJ 07703
  SMOKE/Obscurants, ATTN: DRCPM-SMK, Aberdeen Proving Ground, MD 21005
SUPERINTENDENT:
  US Military Academy, ATTN: Technical Library, West Point, NY 10996
COPIES FURNISHED:
  DOD
    Defense Advanced Research Projects Agency, 1400 Wilson Blvd., Arlington, VA 22209
    Defense Technical Information Center, Cameron Station, Alexandria, VA 22209
  US Army
    HQ Fort Huachuca, ATTN: Tech Reference Division, Ft. Huachuca, AZ 85613
    US Army Field Office, HQ AFSC/SDOA, Andrews AFB, MD 20331
    US Army Inst for Research in Management Information & Computer Sciences,
      115 O'Keefe Bldg., Georgia Inst of Technology, Atlanta, GA 30332
  US Navy
    Naval Air Systems Command, Code AIR-03-D, 1411 Jefferson Davis Hwy., Arlington,
      VA 22202
    Naval Research Lab, ATTN: Director of Research, Wash, DC 20375
    Office of Naval Research, Code 102, 800 North Quincy Street, Arlington, VA 22217
    HQ US Marine Corps, Code RD-1, Washington, DC 20380
  US Air Force
    Air Force Systems Command, ATTN: Deputy Chief of Staff for Science and Tech,
      Andrews AFB, Wash, DC 20334
    Air Force Office of Scientific Research, ATTN: Tech Director, Bolling AFB,
      Wash, DC 20375
  Lawrence Livermore Lab, Univ of California, Livermore, CA 94550
  Los Alamos National Lab, Box 1663, Los Alamos, NM 87545
  NASA HQS, ATTN: Code LD-2, 400 Maryland Avenue, SW, Wash, DC 20546
  National Science Foundation, 1800 G Street, NW, Wash, DC 20550
  United Nations Library, ATTN: Collection Mgmt. Section, Room L-138C, NY, NY 10017

⑫

PROCEEDINGS

OF THE

1984 ARMY SCIENCE CONFERENCE

UNITED STATES MILITARY ACADEMY

WEST POINT, NEW YORK

19-21 JUNE 1984

VOLUME III

Principal Authors P through Z

DTIC
ELECTE
DEC 6 1984
D

D

TABLE OF CONTENTS

PROCEEDINGS OF THE 1984 ARMY SCIENCE CONFERENCE

Author	Title	Vol	Page
Chou, Lucia W.	See Houston, James R.	I	343
Chubb, Douglas W. J.	A Primitive Driven, Artificial Intelligence Threat Model	I	117
Clark, Flynn	See Weiss, Richard A.	IV	283
Cocco, Nick	See Walker, J. F.	V	275
Coleman, Amos J. Howard, Peter D.	Composite Materials for Electrocatalysts	I	131
Cook, Charles F.	See Iafrate, Gerald J.	II	1
Cook, Dennis R. Koenig, Francois J.	Captive-Carry Chaff Experiment	V	31
Corbin, Normand D. Resetar, Theresa M. McCauley, James W. Moon, Kenneth A.	Energy-less Manufacturing of Advanced Ceramics by SHS	I	143
Corliss, Lloyd D.	See Blanken, Chris L.	I	39
Costanzo, Daniel J.	Requirements and Specifications for Cartographic Video Discs	I	155
Cote, Paul J. Meisel, Lawrence	A Combined Magnetic and DTA Method for the Determination of Heat Treatment Data for Armament Steels	I	171
Cronin, Terence M.	Symbolic Filtering of Military Messages Using Artificial Intelligence Techniques	I	183
Dalrymple, Joel M.	See Schmaljohn, Connie S.	III	173
Dart, John	See Bradley, Lester	IV	1
Davio, Stephen R.	Medical Defense Against the Potential Biological Warfare Agent Saxitoxin	I	197
Daws, Robert N., Jr. Keesee, Robin L. Marcus, Arthur Hartel, Christine R. Arabian, Jane M.	Stinger Reverse Engineering Analysis	V	45
Dean, Arnold M., Jr.	Mine Detection in Cold Regions Using Non-Sinusoidal Radar	I	207

A NEW MULTIFUNCTION ACOUSTO-OPTIC SIGNAL PROCESSOR   (U)

JOHN M. PELLEGRINO, DR., NORMAN J. BERG, DR.,
MICHAEL W. CASSEDAY, MR., and *ANDREE N. FILIPOV, MR.
HARRY DIAMOND LABORATORIES
ADELPHI, MARYLAND, 20783.

I. Introduction

An acousto-optic architecture has been devised to simultaneously
obtain time integration correlation and high-speed power spectrum
analysis in a single, compact package. As can be seen in figure 1, this
processor incorporates two TeO_2 shear-wave modulators in a two-level
optical system to obtain two correlator outputs (lower level, figure
1(a)) and two spectrum analyzer outputs (upper level, figure 1(b)).
Information from these sources is sent to a controlling microprocessor,
which analyzes the received signals and reports results to the system
operator. In the present version, the optical system covers a 2 x 3
foot area. This paper discusses considerations involved in the design
and construction of such a processor and presents the operating
parameters and features of the system. Included are correlation results
for a 2-MHz RF chirp input signal and results of multi-signal reception
by the spectrum analyzer section of the processor. Finally, a parallel
combining scheme is proposed to extend the instantaneous bandwidth of
systems in which such a processor may be incorporated.

II. Theory Of Operation

The acousto-optic interaction has been studied extensively and
reports concerning its use in signal convolvers, correlators, and
spectrum analyzers are available.[1-4] Hence only results pertinent to
the operation of this processor will be considered here.

When the incidence angle of the light illuminating the sound wave
in an acousto-optic modulator is set to the Bragg angle
$\theta_B = \sin^{-1}\left(\lambda/(2\Lambda)\right)$ (where λ is the incident light wavelength and Λ the
acoustic wavelength), constructive interference occurs for only a single
first-order diffraction. The intensity of the diffracted light I_D when
operating in this Bragg mode can be shown to be

(a)
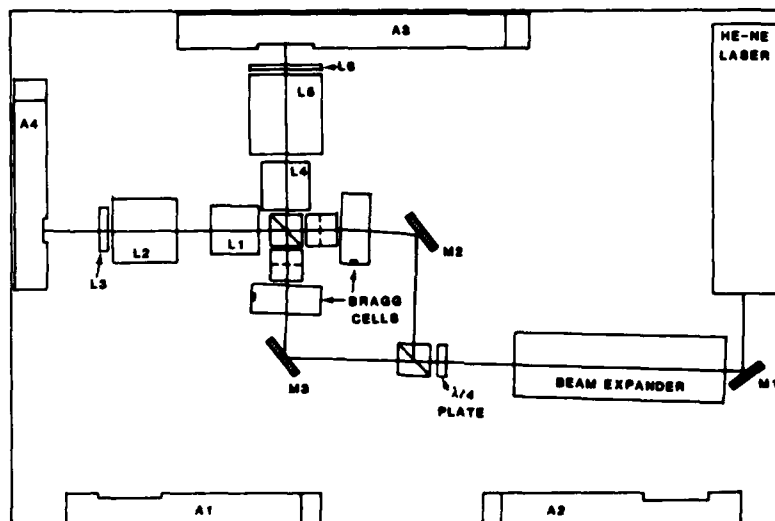
(b)
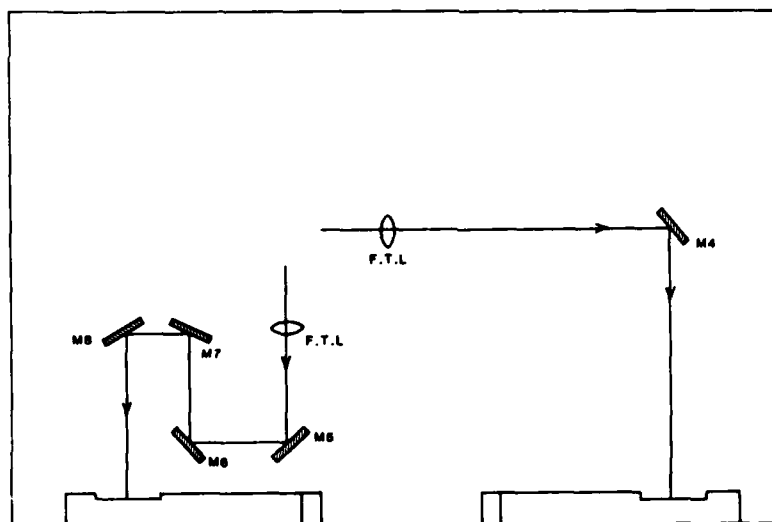
Figure 1. Schematic of the optical system for the multifunction processor: (a) correlator section (lower optical level), and (b) spectrum analyzer section (upper optical level).

$$I_D = I_o \sin^2 \left\{ \frac{\pi}{\lambda} \left(M_2 \frac{W}{2\Lambda} P_a \right)^{1/2} \right\} , \tag{1}$$

where I_o is the incident light intensity; W, the width of the acoustic phase fronts; P_a, the acoustic power; and M_2, an acousto-optic figure of merit of the delay line material. For sufficiently small P_a, the diffracted light intensity for well-collimated light may be considered linear with acoustic power.

When a bandpass signal, $A(t)\cos\omega(t + \phi)$, is applied to the modulator, it generates a sound wave, $S(t,z)$, which propagates through the Bragg cell. This wave may be described as

$$S(t, z) = A'(T - z/\upsilon) \cos\omega(t - z/\upsilon + \phi), \tag{2}$$

where υ is the acoustic propagation velocity, and z is the distance along the Bragg cell. The diffracted light may then be represented by

$$L_D(t, z) = L_o A''(t - z/v) \cos\left(\omega_\ell t + \omega \left(t - \frac{z}{\upsilon} + \phi \right) + \frac{z \sin \theta_B}{\lambda} \right), \tag{3}$$

where ω_ℓ and L_o are the incident light frequency and the amplitude, respectively. The diffracted light $L_D(t, z)$ is seen to contain all the signal information between the time t and $t - Z/\upsilon$, where Z is the illuminated length of the Bragg cell, and to exit the cell at the Bragg angle, which may be considered to be proportional to the frequency of the signal for the light and signal frequencies commonly used in these devices. Hence a spectrum analyzer may be constructed simply by placing a Fourier transform lens behind the modulator to collect and focus the diffracted light onto an array placed at the focal plane of the lens, position along the array corresponding to input signal frequency.

Diffracted light from each of the two modulators (figure 1(a)) may be combined and filtered to remove any undiffracted light to perform correlation. Diffracted light from each cell i (i = 1, 2) is described by

$$L_1(t, x_1) \approx S_1 \left(t - \frac{x_1}{\upsilon} \right) \cos \left\{ \omega_\ell t + \omega \left(t - \frac{x_1}{\upsilon} \right) + \frac{x_1 \sin \theta_B}{\lambda} \right\}, \tag{4}$$

for $\theta_B = \sin^{-1}\left((\Lambda - \Lambda_o)/\lambda \right)$, where Λ_o is the acoustic wavelength at the Bragg-cell design center frequency, ω_o. Beams $L_1(t, x_1)$ and $L_2(t, x_2)$ are combined using a cube beam-splitter, filtered, and so illuminate the time-integrating square-law photodetector arrays A_3 and A_4. The output

3

of these arrays may then be shown to be[5]

$$V(z) \sim \cos\left(\frac{2}{M}(\omega - \omega_o)\tau\right) \int_T S_1(t')S_2(t' + \frac{2\tau}{M} - \tau_o)\,dt, \qquad (5)$$

where $\tau = z/\upsilon$, $\tau_o = Y/\upsilon$, $t' = t - \tau$, M is the magnification factor of the optical filtering system, and Y, the illuminated length of Bragg cell 2. Equation 5 is thus seen to be the correlation of the bandpass signals $S_1(t)\cos\omega t$ and $S_2(t)\cos\omega t$ offset by the frequency ω_o and in a compressed, shifted time frame. Given that one uses a single laser source and hence forms a coherent optical system, the output is a detected interference pattern of the diffracted light from the two modulators. As such, this correlator is generically similar to the two-beam surface-acoustic-wave (SAW) time integrating correlator previously reported.[6,7]

III. System Design

The size and performance of the processor are affected by both the type of Bragg cell used and the length and diode spacing of the available photodiode arrays which detect the spatial light distributions. If a 50-µs Bragg-cell delay length is used to insure 25-kHz frequency resolution, the physical length (and illuminated aperture) of a $LiNbO_3$ SAW cell would have to be 17.5 cm, while a $Bi_{12}GeO_{20}$ SAW Bragg cell would be 8.7 cm, and a TeO_2 slow-shear bulk-acoustic-wave Bragg cell would be 3.1 cm. Clearly the TeO_2 cell results in the smallest processor as far as the input optics (laser, beam expander, etc.) are concerned. In addition, the greater angular beam-deflection versus frequency of this cell results in much shorter focal length (thus more compact) Fourier-transform lens system to achieve the spot travel required by a given diode-array length for spectrum analysis. The TeO_2 cell offers another advantage in its high diffraction efficiency, which for a single input signal may reach or even exceed 60 percent depending upon the particular cell.

This compactness is achieved with several disadvantages. The shear-wave cell operates maximally with incoming light circularly polarized. As the crystal is optically active, the modulated light is elliptically polarized. Since a large portion of the light exiting each Bragg cell is split off for use in the spectrum analyzer section of the processor, care must be taken to ensure that the light from each cell reaching the correlator arrays is of matched polarization so that optical interference can be observed. This entails appropriate design and custom coatings for beam-splitter and combiner cubes.

4

Phase distortion in the modulated laser light is also seen to be a problem, in contrast to the virtually distortion-free SAW devices. With Isomet OPT-1 Bragg cells, diffraction performance across the light aperture is good in the direction corresponding to that of sound propagation in the device. Perpendicular to that direction, a variation in diffraction intensity can be observed, which affects the depth of the interference fringes when two beams similarly perturbed are combined. The cause of this phenomenon, although uncertain, may be due to defects or domains in the TeO_2 crystal. Performance under these circumstances may be improved by selective aperturing of the device and/or modulated light which is to be focussed onto the detector array.

A more serious disadvantage is that the spatial-frequency variation versus signal-frequency change at the correlator detector array is greatly increased. From equation 5, the array signal variation with z, the distance along the array, is seen to include a term

$$V(z) \sim \cos[\frac{2}{M}(\omega-\omega_o)\frac{z}{\upsilon}].$$ (6)

For the 617-m/s sound velocity in TeO_2 this yields a spatial frequency of about 3.3×10^4 cycles/m for $M = 1$ and $|\omega - \omega_o| = 2\pi \times 10.3$ MHz thus a maximum bandwidth of 20.6 MHz can be sampled by a commercially available detector array with 15-μm diode spacing. For the same array diode spacing, a correlator using $LiNbO_3$ SAW cells would have a usable bandwidth of 114 MHz.

The usable correlation bandwidth with TeO_2 cells can be increased by magnifying the diffracted beams with a spatial filter system. Using a 2 : 1 magnification ratio ($M = 2$), a commercially available 62-mm long, 4096-diode array allows a 25-MHz signal bandwidth with about 3.3 samples per spatial cycle and displays the full time aperture of the cells. This bandwidth is achieved in commercially available TeO_2 cells with 50-μs delay apertures, providing a performance-matched system.

The second array used for correlation detection is illuminated by only the center section of the diffracted light beams. An aperture stop blocks light from all but the center 4 μs of the Bragg cells, and the spatial extent of this light is magnified by a factor of 10 in the second spatial filter system. The expanded output illuminates a 25.6-mm long, 1024-diode high-dynamic-range detector array that provides high resolution time-difference-of-arrival (TDOA) information. The minimum correlation width expected with 30-MHz bandwidth signals is approximately 67 ns. In the compressed t space of the correlator, this corresponds to a spatial extent of about 0.2 mm after magnification

M = 10. This would be sampled by 8 diodes, allowing TDOA interpolation to within 10 ns.

IV. Spectrum Analysis

The spectrum analyzer section of the processor is illustrated in figure 1(b). Collimated, coherent, circularly polarized laser light (5-mW Helium-Neon) is split using a 50/50 nonpolarizing cube beam-splitter; each beam from this splitter illuminates one of the modulators, incident at the Bragg angle for the cells' center frequency of 45 MHz. The diffracted light in a each cell is split into two parts, one of which is sent vertically to the optical level where, after passage through a 120-mm Fourier transform lens, it is focussed onto a high speed photodiode detector array.

Arrays A_1 and A_2 were chosen for their ultra-fast readout capabilities: each array is read out serially using a combination serial-parallel scheme. The photodiode array has 1024 elements

SPECTRUM ANALYZER DUAL SIGNAL RECEPTION

100 KHz SEPARATION   80 KHz SEPARATION

65 KHz SEPARATION   50 KHz SEPARATION
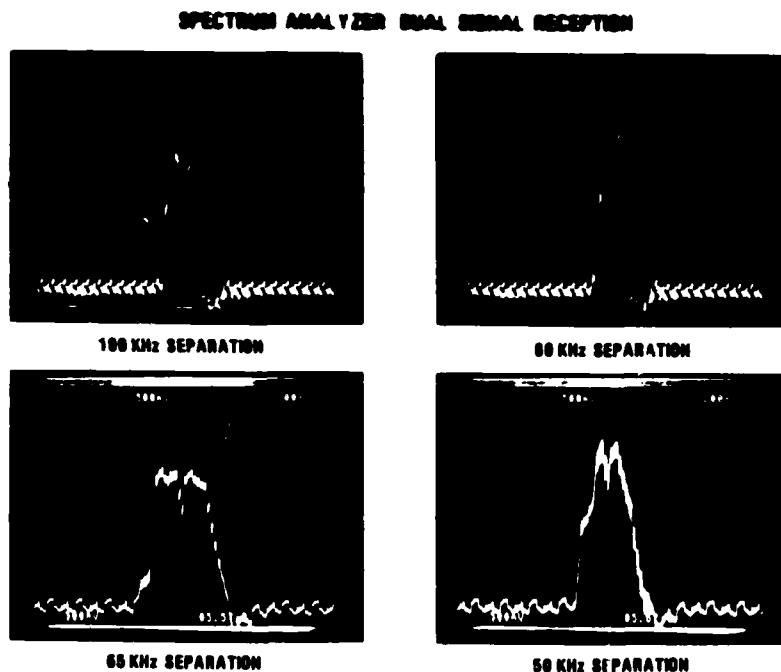
Figure 2. Output of the spectrum analyzer for two simultaneous cw RF input signals with various frequency spacing, Δf.

6

separated into eight segments; these segments are read out in parallel while the diodes within each segment are read out serially. This high-speed capability is necessary for time-of-reception determination of signal frequency changes. Every 15 µs, an update report is sent to the system controller indicating the arrays' status. One disadvantage of these arrays is the high light intensity necessary to trigger them; the beam-splitter cubes were designed to send 90 percent of the incident light to spectrum analysis processing. Each array covers the full 25 MHz of system processing bandwidth. When the signals received from each of two directional antennas are used as separate inputs to each Bragg cell, direction of arrival information may then be derived via amplitude comparison of these two arrays.

CORRELATOR: 2 MHz CHIRP

BW= 2 MHz
Δf = 64 KHz
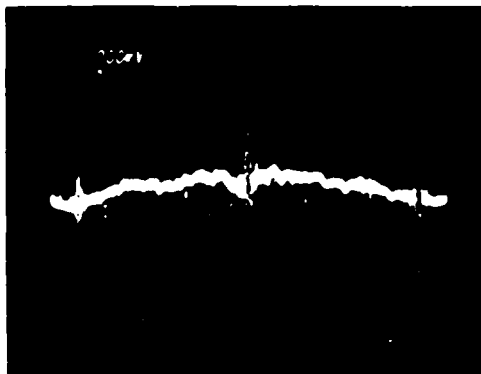
BW=2 MHz
Δf=128 KHz
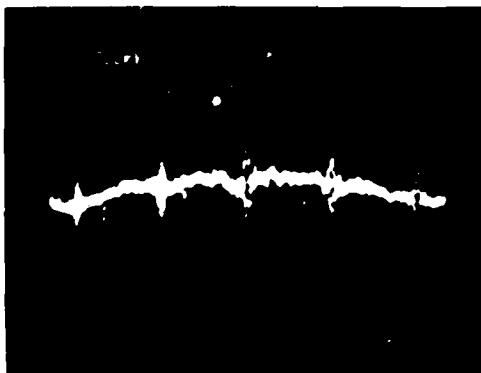
Figure 3. Correlator output for a 2-MHz bandwidth step-wise chirp input to both Bragg cells. Shown are foldover peaks for 64- and 128-kHz frequency spacings.

With this configuration, the optical spot size for a single-frequency RF signal is about 60 μm, covering three pixels on the detector. Hence an individual signal can be resolved to within 25 kHz. As shown in figure 2, signals received simultaneously can be distinguished when they are as little as 50 kHz apart. Signals received sequentially in time, however, can be distinguished with frequency differences as little as 25 kHz. Single signals can be distinguished with this resolution over a dynamic range of approximately 25 dB. This figure is achieved when background electronic-array-noise and signal due to spurious light reflections are subtracted from the signal readings.

V. Time Integrating Correlation

The 10 percent of remaining diffracted light is combined in a beam-splitter/combiner cube and sent to the two time-integrating photodiode detector arrays. Notice in figure 1(a) that the optical path has been

CORRELATOR: EXPANDED CENTER
2MHz CHIRP

FULL APERTURE   CENTER

fc = 40 MHz
BW = 2 MHz
Δf = 64 KHz
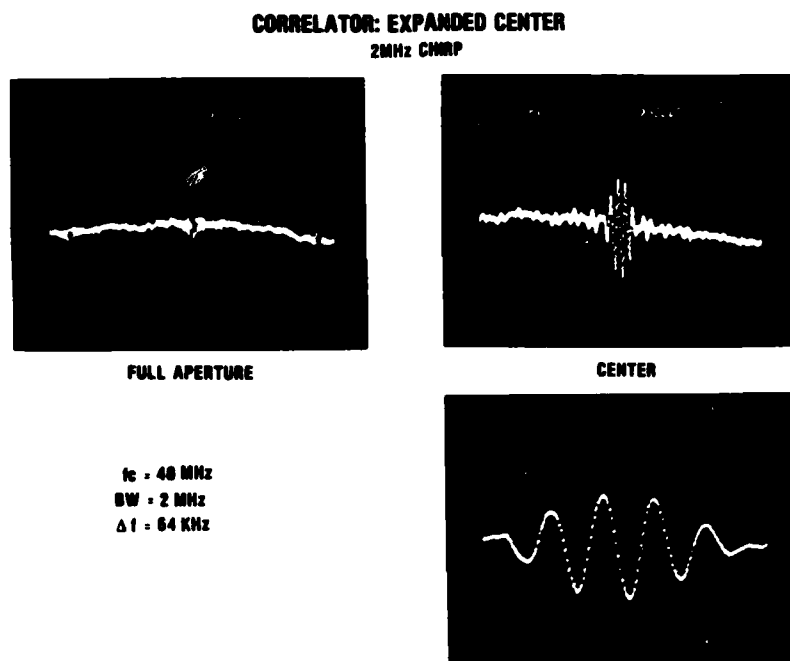
Figure 4. Output of the M = 10 correlator section for the same 2-MHz bandwidth chirp used in figure 3, with Δf = 64 kHz. Shown is full aperture and the center 4 μs of correlation space. Center frequency of the chirp f_c is 40 MHz.

8

kept as compact as possible. This was done to minimize the walk-off
which occurs when interfering beams are combined at greater angles of
intersection, corresponding to maximum deviation from the device center
frequency. In this configuration, walk-off has been minimized such that
the two beams overlap by more than 80 percent, assuring correlation of
nearly the full time aperture available.

Figure 3 shows the output of array A_3, which is set to observe the
full 50 µs of delay aperture and integrate over a 30-ms time period.
The input to each cell consisted of a 2-MHz step-wise chirp, where
increasing frequencies are stepped through sequentially. It is an
interesting correlation property of these signals that when the system
resolution of 25 kHz is finer than the step size Δf, foldover peaks can
be observed to either side of the main correlation. The number of these
peaks is proportional to Δf. The output of array A_4, which is set to
sample the center 4 µs of delay aperture, is shown in figure 4 for the
same input signal used for figure 3. The position of this correlation
can be easily determined and the information used to obtain TDOA and,
hence, direction of arrival.

VI. Bandwidth Enhancement

The acousto-optic signal processor that has been described is
limited by practical material considerations, to about 25 MHz
instantaneous bandwidth. Figure 5 illustrates a parallel combining
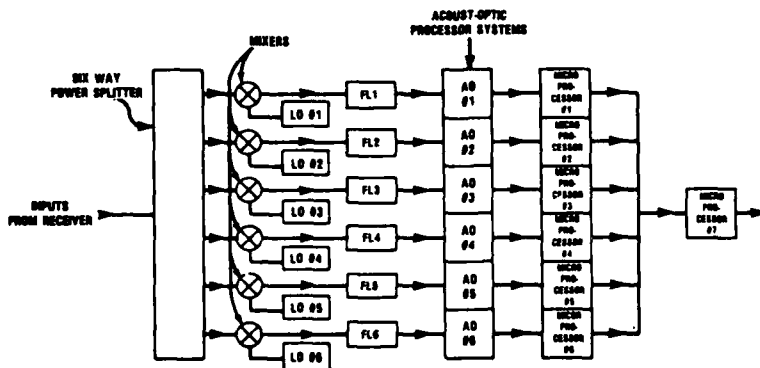
Figure 5. Parallel combining scheme for processor bandwidth
enhancement.

scheme for increasing the instantaneous bandwidth of a receiver system employing these processors, without sacrifice of resolution or other processor performance.

In the figure six identical processors are shown, each taking its input signals from the corresponding IF outputs of two 150-MHz total-instantaneous-bandwidth (input bandwidth) channelized receivers. Each acousto-optic processor output is digitally post-detection processed as a stand-alone system. The outputs from the six post-detection processors are then further analyzed by a system which digitally restores the frequency offset information to each individual output and examines the combined signal. The result is a 150-MHz bandwidth spectrum analyzer of increased time-bandwidth product with 25-kHz frequency resolution and 15-us time-of-reception resolution, combined with a 150-MHz bandwidth time-integrating correlator able to detect signals and resolve TDOA to better than 10 ns (for maximum bandwidth signals).

VII. Conclusion

An acousto-optic processor that combines the features of high-speed spectrum analysis and time-integrating correlation in a fairly compact optical package has been demonstrated. A new version in more compact form is in the planning stages; this system would be rack-mountable and packaged for RF shielding. Operating parameters will be identical. These processors may be combined for extended bandwidth operation or used as convenient stand-alone operating systems.

References

1). R. Adler, "Interaction Between Light and Sound," IEEE Spectrum, 4, 5, (1967), pp 42-54, and references therein.

2). I. C. Chang, "Acoustooptic Devices and Applications," IEEE Trans. Sonics Ultrason., 50-23, (1976), p 2.

3). W. T. Rhodes, "Acousto-Optic Signal Processing: Convolution and Correlation," Proc. IEEE, 69, 1, (1981), p. 65.

4). T. M. Turpin, "Spectrum Analysis Using Optical Processing," Proc. IEEE, 69, 1, (1981), p 79.

5). N. J. Berg and J. M. Lee, ed., Acousto-Optic Signal Processing: Theory and Implementation, Marcel Dekker, Inc., New York (1983).

6). M. W. Casseday, N. J. Berg, I. J. Abramovitz, and J. N. Lee, "Wide Band Signal Processing Using the Two-Beam Surface Acoustic Wave Acousto-Optic Time Integrating Correlator," IEEE Trans., MTT-29, 5, (1981).

7). N. J. Berg, M. W. Casseday, I. J. Abramovitz, and J. N. Lee, U.S. Patent #4,326,778, 27 April 1982.

EVALUATION AND INTERPRETATION OF COMPRESSOR INTRA-BLADE FLOW FIELD MEASUREMENTS OBTAINED USING AN ADVANCED LASER ANEMOMETER SYSTEM

MICHAEL J. PIERZGA

U.S. Army Propulsion Laboratory (AVSCOM)
NASA Lewis Research Center
Cleveland , Ohio 44135

INTRODUCTION

Advancements in the aerodynamic design of gas turbine engine fans, compressors and turbines have historically resulted primarily from an empirical approach. Although this approach has led to more efficient, lighter weight engines, major new advancements in component performance will most likely not result from empiricism alone. A fundamental understanding of the complex flow phenomena which occur within advanced turbomachines is essential if significant advancements in performance are desired. As the cost of building and testing gas turbine engine components increases so does the need to reduce the number of redesigns necessary to achieve the desired performance.

In addressing this need, the NASA Lewis Research Center, in cooperation with the Army Propulsion Laboratory, has undertaken a program designed to increase the basic understanding of the complex internal flows of gas turbine fans and compressors. The program relies primarily on the use of non-intrusive optical techniques to evaluate the intra-blade flow field of transonic compressor blading. Concurrently, advanced computational techniques are being developed which incorporate the three-dimensional flow field effects of this class of turbomachine.

Previous researchers have presented comparisons between 3-D flow calculations and internal flow field measurements [1] and [2]. They have generally found that inviscid codes model the inlet to transonic rotors quite well but tend to over predict the deceleration through the rotor. This is usually attributed to lack of the streamline displacement effects due to boundary layer growth. Those who have added boundary layer calculations to 3-D analyses [3] and [4] have found that the flow field is substantially altered by addition of the displacement thickness to the solid surfaces. The technique used in [3] included simple models for shock-boundary layer interaction, laminar-turbulent transition, and

13

separation effects. The boundary layer calculations were performed only on the blade surfaces and yielded better agreement with experimental data than did the calculations without boundary layer effects. In [4] results were shown for the same rotor analyzed in [1] and demonstrated the extent to which inclusion of a simple boundary layer model on the blades and endwalls could alter the flow field in a transonic rotor.

The focus of the present work is the investigation of a detailed set of optical measurements of an advanced transonic, axial-flow compressor flow field obtained under the current NASA/Army research program. These data, coupled with the results of an inviscid, three-dimensional flow field calculation, provide the basis on which the internal flow features and trends are interpreted herein. A more complete discussion of the present work can be found in reference [5]. Strazisar [6] provides a more fundamental overview of the rotor flow physics including the three-dimensional nature of the shock fronts based on these same data.

COMPRESSOR ROTOR

The research vehicle used in the present work was the first stage rotor of a NASA Lewis designed two stage fan [7]. This first stage rotor was designed as a 22 blade, low aspect ratio (1.56), damperless replacement for the original 43 blade, high aspect ratio (2.94), dampered rotor [8]. The new blading met or exceeded its design total pressure ratio and adiabatic efficiency goals of 1.629 and 0.896, respectively, when tested in the two stage configuration. The inlet relative Mach number at the rotor tip was 1.38 at the design tip speed of 428.9 m/s and design mass flow rate of 33.25 kg/s.

The rotor was tested in the present work without inlet guide vanes or downstream stators so that blade row interactions would not be present. Figure 1 shows the measured rotor performance characteristic for a constant design speed condition. Also shown in the figure, for reference purposes, are data from the redesigned two stage fan investigation [7]. Herein, all future reference to rotor performance will deal solely with the isolated rotor data.

The entire experimental research program conducted with the isolated rotor configuration of this low aspect ratio fan covered a range of data points from max flow (MF) to near stall (NS); however, the most detailed information was obtained at the peak efficiency (PE) and near stall points. Most of the comparisons between experiment and computation will be restricted to these two flow points.

INSTRUMENTATION

A fringe laser anemometer system (LA) was the principal instrument

used during the experimental evaluation program. This instrument is a single channel, dual beam system with an on-axis backscatter collection scheme and has been previously described in detail [9] and [10]. Data aquisition and storage was accomplished via a dedicated mini-computer while post processing and graphical output was handled by a large central computer. Access to the compressor flow path was through a 3 mm thick glass window which extended far enough foreward and aft of the test rotor to enable free-stream and blade wake measurements. The window curvature closely conformed to the rotor outer flow path thereby minimizing disturbances to the tip region flow. Fluorescent seed particles, whose nominal diameter was 1-1.4 microns, were spray atomized into the flow stream through a 6 mm diameter tube located over 60 cm upstream of the rotor leading edge. In addition to LA measurements, conventional instrumentation was used to record compressor overall pressure and temperature levels. This information was used to calculate the compressor overall performance characteristic as well as to control the on-line operating condition set point.

The laser anemometer survey locations along with the meridional view of the computational grid used by the 3-D analysis code are shown in figure 2. LA survey locations, denoted by solid circles in figure 2, were distributed axially at specific locations based on the axial projection of the local chord length along design streamlines to provide data coverage at 20% intervals upstream, 2.5% intervals around the leading edge, 10% intervals within the blade and five evenly spaced locations between the rotor trailing edge and the position of the stage configuration stator leading edge.

NUMERICAL ANALYSIS

The numerical analysis of the rotor was conducted using a three-dimensional computational procedure developed by Denton [11] . This procedure was based upon a finite volume technique. The code was modified [4] to include the effects of boundary layer displacement using an injected mass technique which allowed flow to pass through the walls in proportion to the rate of growth of the boundary layer displacement thickness. The boundary layers were calculated as if the grid lines in the streamwise direction were actual flow field streamlines. Boundary layers were calculated on both the blade surfaces and on the endwalls which were assumed to rotate with the blading. The displacement thickness was computed from the von Karman integral boundary layer equation using a constant shape factor (1.5) and constant skin friction coefficient (.005). The boundary layer displacement thickness was updated every 50 timesteps with 50 percent of the new displacement thickness added to 50 percent of the previous value. Fractional updating of the boundary layer was necessary since full correction tended to

15

Figure 2. Meridional plane computational grid including LA survey locations.

Figure 1. Measured rotor performance at design speed.

produce instability.

The upstream boundary was located 1.5 times the tip axial chord upstream for the tip and 0.7 times the hub axial chord upstream for the hub. At the upstream boundary the total pressure, total temperature, absolute tangential velocity, and meridional flow angle were held constant at each spanwise grid line and assumed to be uniform in the tangential direction. The only downstream condition specified was the static pressure at the hub and it was assumed to be constant in the tangential direction. The static pressure gradient from hub to tip at the downstream boundary was calculated assuming no curvature of the streamlines in the hub-to-shroud plane which was appropriate for this case based upon the flow path geometry. The downstream boundary was located twice the tip axial chord downstream at the tip and one hub axial chord downstream at the hub.

All of the solutions presented were obtained with a computational grid of 21 points blade-to-blade, 95 points in the streamwise direction with 50 points from blade leading to trailing edge and 11 points from hub-to-shroud. Streamwise grid spacing was constant on the blading but was expanded upstream and downstream in order to isolate the boundaries from the leading and trailing edges with as few points as necessary in an attempt to capture the correct upstream wave strength and to limit the effect of a uniform circumferential pressure at the downstream boundary. In all cases multi-grid (7x12x3) with a non-uniform timestep was used.

Convergence was assumed to occur when all the following conditions were met:

(1) The maximum change in axial velocity anywhere in the flow field divided by the root-mean-square of all the absolute velocities in the flow field did not exceed .01 percent.
(2) Mass flow at each streamwise grid line agreed with the inlet mass flow to within 0.3 percent.
(3) Inlet mass flow did not change by more than 0.3 percent over several hundred timesteps.
(4) Mach numbers at four selected positions on the rotor suction surface did not change by more than 0.02 over several hundred timesteps.

For the computations with the boundary layers included, timesteps to convergence ranged from 2486 for the peak efficiency case where the location of the shock is very sensitive to even small mass flow changes to 1300 at the maximum flow condition. Computation times on a Cray-1 computer were 17 minutes and 9 minutes, respectively. Computation time on an IBM 370-3033 would be approximately 10 times as long. No appreciable increase in time resulted from use of the boundary layer

17

calculations. The solution actually converged somewhat better with the boundary layer calculations than without.

RESULTS AND DISCUSSIONS

Introduction

The data presentation initially focuses on the use of contour maps of relative Mach number to show changes in flow conditions as the rotor exit pressure was changed. Comparisons are made between the analysis code and the results obtained with the laser anemometer. Results are also presented in terms of overall pressure rise as measured and predicted. Plots of relative Mach number at various flow rates and percent spans are shown to give a more detailed picture of the flow in a quantitative sense. In each line plot the analysis is compared to the LA data at comparable percent gap locations. Finally, some applications of the results which might have significance in the design of future compressor blading is addressed.

Mass flow rates measured in experimental facilities and calculated numerically have a degree of uncertainty associated with them such that comparison of absolute magnitudes may not match well. To compensate for this fact, all of the comparisons are made at mass flow rates that have been non-dimensionalized with the maximum flow rate measured from the experiment or calculated by the analysis code as appropriate. Although this appraoch was taken, the agreement between the predicted maximum flow and the experimental measurement was excellent for both the runs without and with boundary layer calculations. The maximum flow predicted without boundary layers was 1.2 percent higher than the measured flow and that predicted with boundary layers was only 0.06 percent higher.

Relative Mach Number Contours

Thirty percent span. Figure 3 is a composite showing the results obtained with the laser anemometer and analysis code at 30% span for the three flow conditions; max flow, peak efficiency, and the near stall point. The shock locations are shown for each flow rate in order to provide a clearer picture of the flow in the passage. Locations of the wave systems and shocks were determined by inspection of relative Mach number and relative flow angle data with respect to the streamwise direction and designating the starting point of the flow deceleration as the shock front. This was done at as many pitchwise locations as needed for the desired resolution. Experimental LA data must be interpreted in this manner due to possible seed particle lag once a high velocity gradient region is encountered. The numerical calculation actually produced shocks that were smeared over several grid points and thus the actual location of the shock is between the begining and ending of the
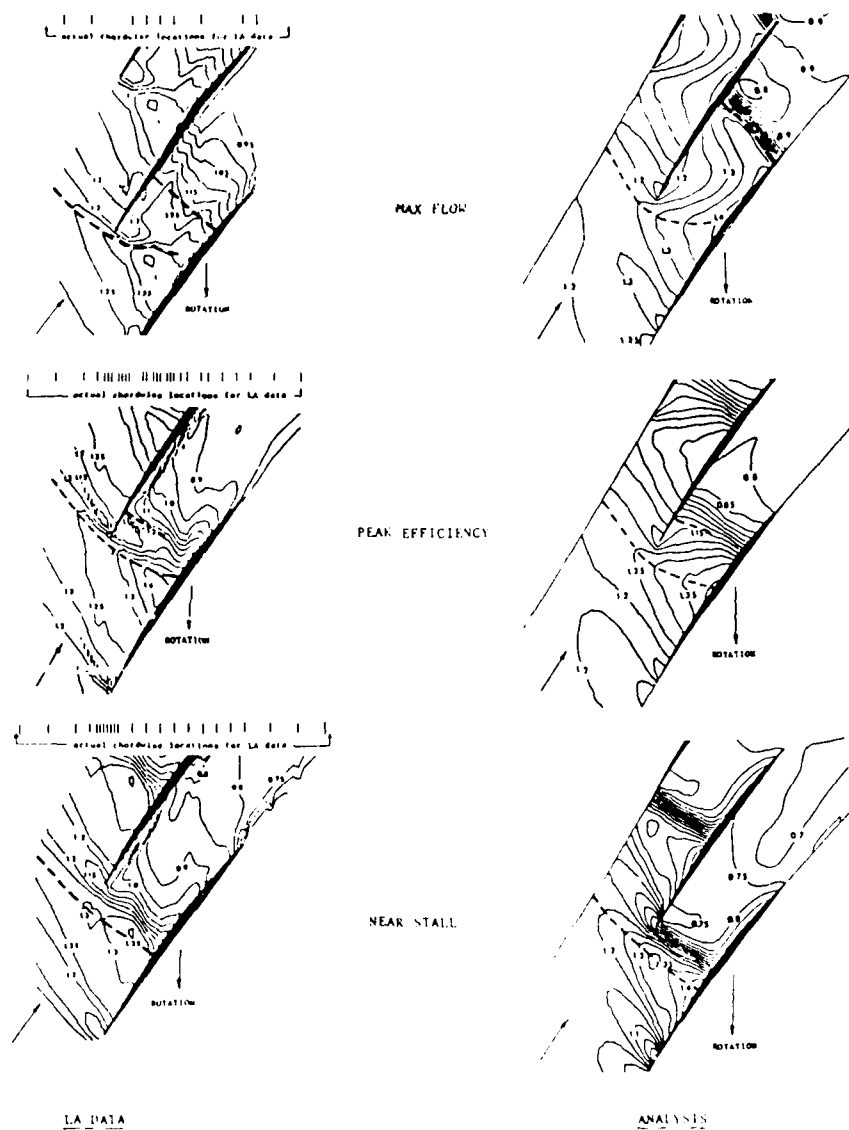
Figure 3. Contour maps of relative Mach number at
30% span from the tip: Comparison between LA
survey data and analysis results.
-------- Indicates Shock Location

deceleration. The grid spacing used herein was small and the shocks were assumed to be located at the beginning of the deceleration.

The LA data in figure 3 shows the progression from an oblique shock at max flow with supersonic flow in the aft portion of the blading, to a near normal shock followed by a second shock at peak efficiency, and finally, a single normal shock ahead of the blade leading edge at the near stall point. The peak Mach number varies from 1.4 ahead of the shock at max flow and peak efficiency to 1.35 ahead of the shock at near stall. Also shown in the figure are the results obtained with the 3-D analysis code. The major difference between the experiment and analysis at max flow is the location of the second shock at the exit of the passage. The peak Mach numbers upstream of the first shock are approximately equal. Although, the location of the second shock is sensitive to backpressure at the max flow condition, a 7% increase in backpressure was required in order to move the computed shock to the approximate position of the measured second shock. This is a large change in pressure and it is more probable that the difference in location is a result of the constant shape factor used in the boundary layer calculations. Also, LA data in the aft portion of the rotor for the max flow condition was only taken at the 40, 60, 90 and 100 percent chord locations so that good shock definition in this region was very difficult.

The agreement at the peak efficiency point between the data is much better than for the max flow case. The front shock is accurately located by the code and there is a rear shock located in approximately the same position as the measured rear shock. The second shock is shown extending from the pressure surface to about mid-pitch. At points closer to the suction surface, interpreting the location of the shock was more difficult since an acceleration of the flow or a constant Mach number region was not found in the streamwise plots. However, indications of a second shock were found near the suction surface by noting changes in the Mach number gradient in the streamwise direction. This was true for both the analysis and the LA data. The computed Mach number in the exit region of the blade is lower than the measured Mach number (0.8 compared to 0.9). Peak Mach number upstream of the shock is 1.4 for both the data and the analysis. The agreement between analysis and data at the near stall point is good in terms of shock location and peak Mach number. The computed Mach number at blade exit is under predicted (0.7 versus 0.8), as it was with the previous cases.

Near tip and near hub sections. Figure 4 shows the peak efficiency and near stall flow points at the 10% span location (figure 4a) where the inlet relative Mach number is supersonic and at the 70% span location (figure 4b) where the inlet relative Mach number is subsonic. Results

LA DATA
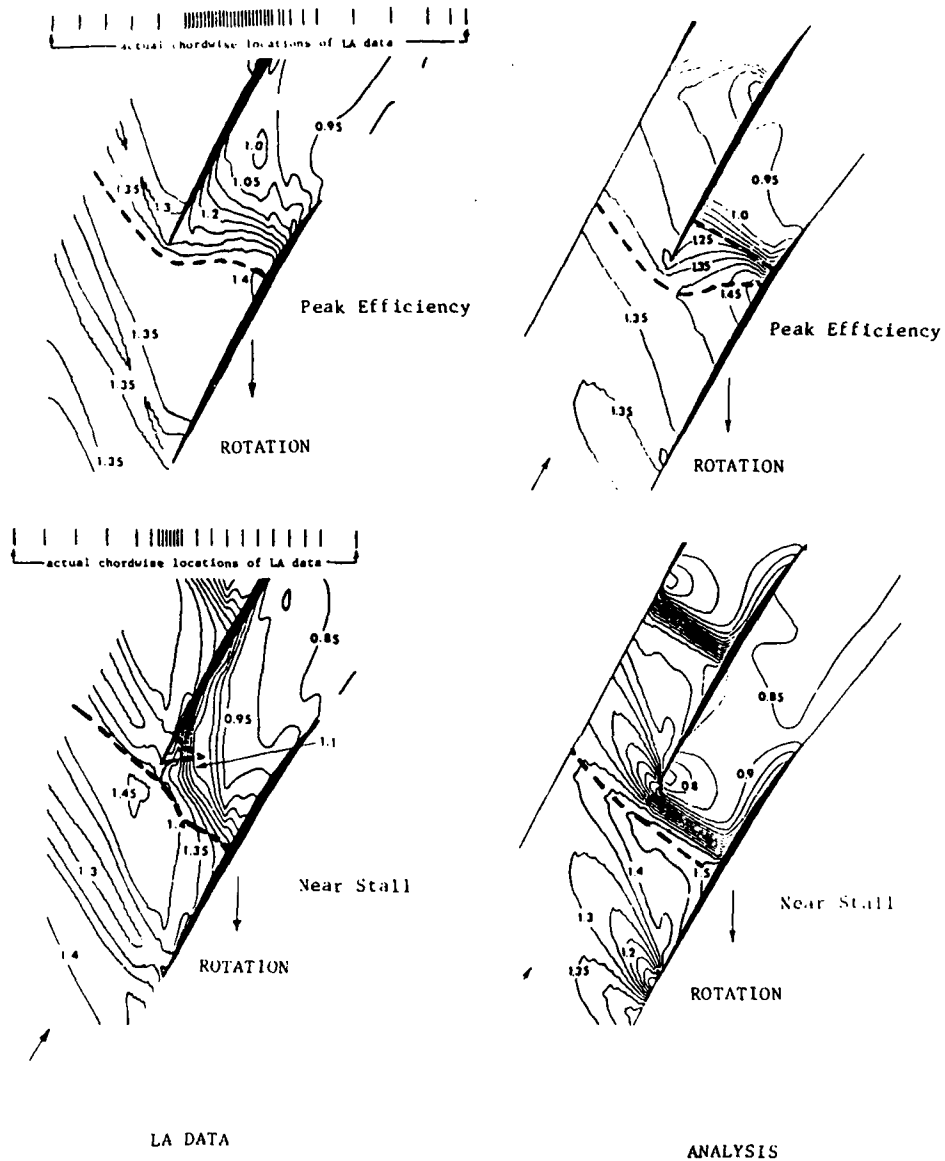
ANALYSIS

Figure 4a. Contour maps of relative Mach number
at 10% span from the tip: Comparison between LA
survey data and analysis results.
-------- Indicates Shock Location
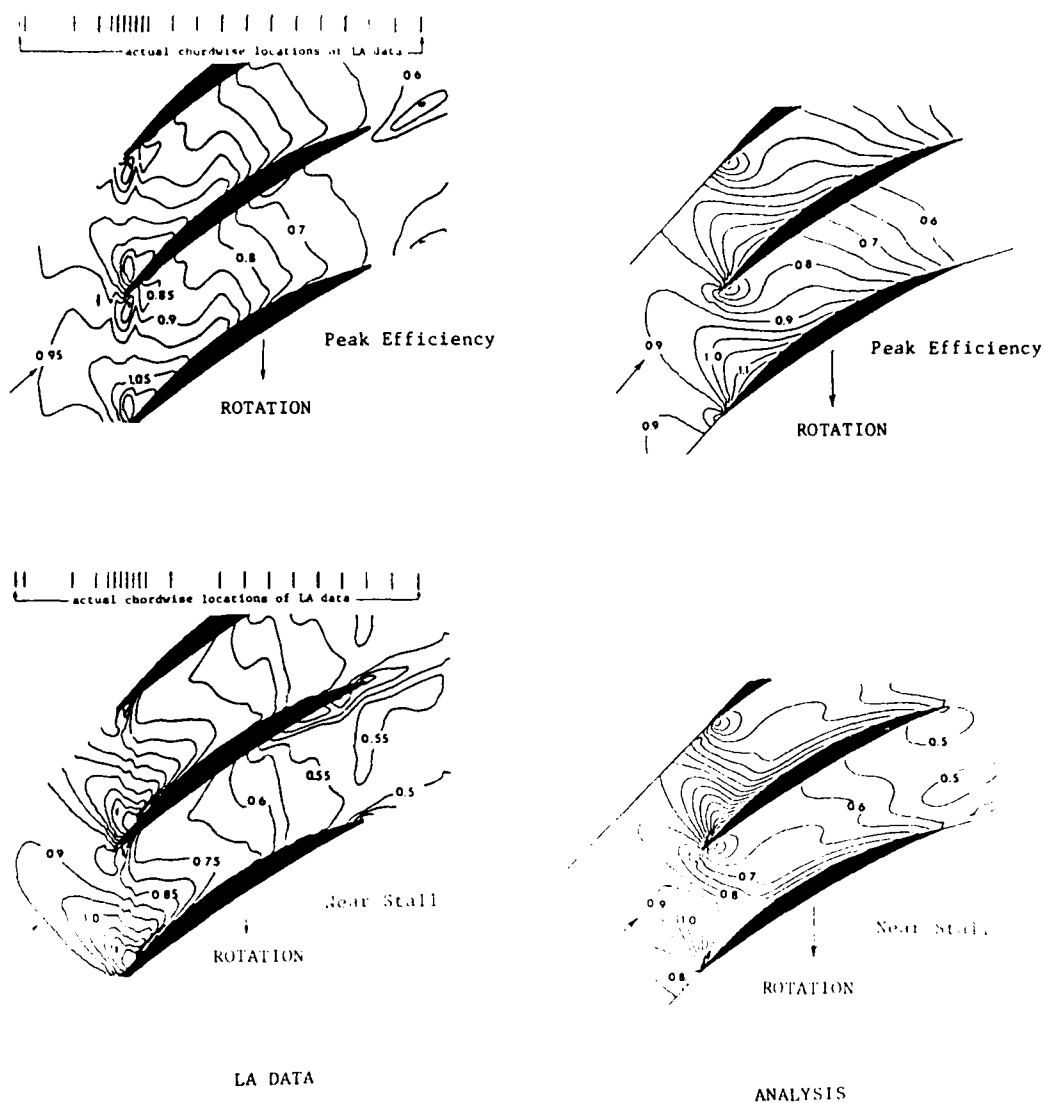
Figure 4b. Contour maps of relative Mach number at 70% span from the tip: Comparison between LA survey data and analysis results.

are shown for both the experiment and the analysis. The peak efficiency
LA data at 10% span in figure 4a shows a front shock that is nearly
normal around the leading edge but is oblique over most of the passage
gap. Inspection of the streamwise and pitchwise LA data indicated that a
second shock might exist. However, the data was not conclusive and thus
a second shock is not indicated in the figure. A second shock was
predicted by the 3-D analysis code at the approximate location suggested
by the LA data. The high values of relative Mach number after the normal
portion of the shock have been investigated by Strazisar [6] and appear
to be due to the 3-D shock inclination in the hub-to-shroud plane.

At the near stall point the second shock has disappeared from the
analysis and only a normal shock standing in front of the blade remains.
A second shock system on the pressure surface around the leading edge is
indicated by the LA data. This feature appears to be necessary to guide
the high incidence angle flow around the leading edge so that it can
adjust to the blade surface angles. This assumption was supported by
computing particle trajectories through this proposed system based on the
wave angles and the measured inlet Mach number. Final pressure surface
flow angles after passing through the second shock system were within a
few degrees of the surface values. This phenomenon is not predicted by
the analysis and may be due, at least in part, to tip region flow
effects. More extensive analyses at other operating conditions and
percent span locations with high incidence angles is required before a
full explanation can be presented. Peak Mach numbers before the shock
are, in both cases, approximately 0.1 higher for the analysis than for
the data. Suction surface Mach numbers are generally higher (0.1 to 0.2)
in the analysis since the viscous effects are incorporated only as a mass
addition and not a momentum deficit.

Figure 4b shows maps of constant relative Mach number for peak
efficiency and near stall at the 70% span location for both the LA data
and the 3-D analysis results. Flow field features between the two
operating conditions are very similar and agreement between the LA data
and the numerical results is quite good although the analysis yields a
slightly lower exit Mach number than experimentally verified.

Overall Performance

In figure 5 an attempt has been made to provide a clear picture of
how the 3-D analysis code compares to the experimental data on a global
basis that could be useful in the design phase of a new blade row.
Figure 5 shows the static pressure rise across the rotor at the hub and
tip as a function of non-dimensional mass flow rate. As expected, the
analysis without boundary layers predicted a much higher pressure rise at
both hub and tip than was measured with conventional pressure
instrumentation. Excellent agreement between the measurements and the
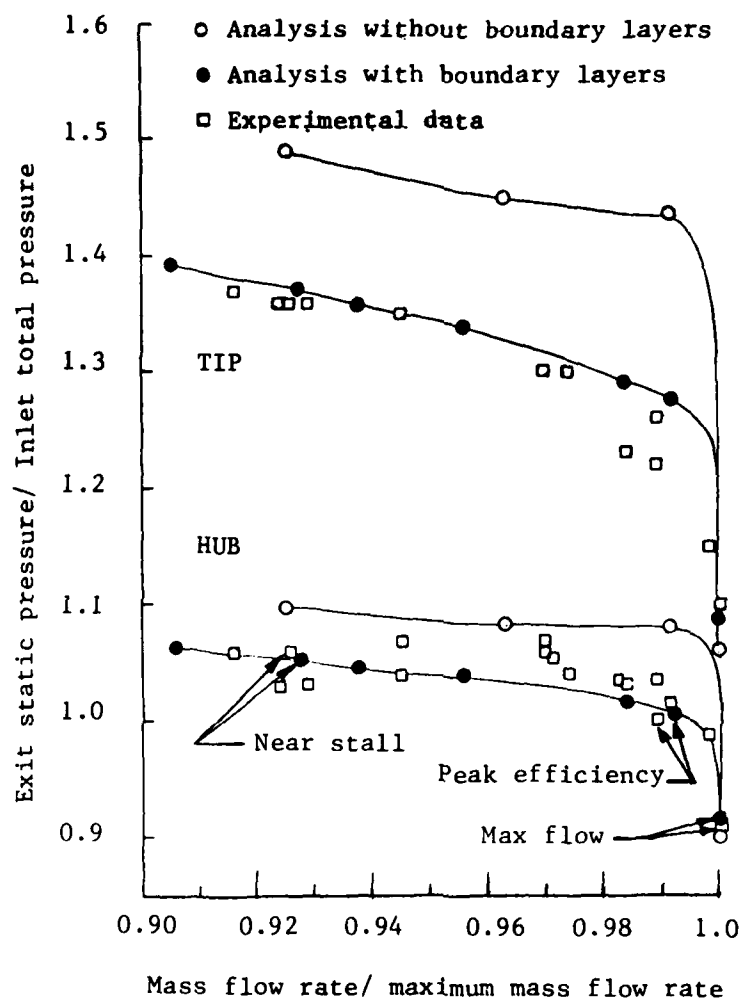
Figure 5. Rotor tip and hub static pressure ratio.

analysis with boundary layer effects is experienced over the entire speed line from max flow to near stall. Unstable operation of the rotor was experimentally observed at a mass flow ratio of about 0.89, while at a mass flow ratio of 0.906 the analysis was stable. The near stall numerical solution was certainly aided by the restriction of a constant shape factor for the boundary layer calculations.

Near-Blade Surface Comparisons

Generally, an important task for the blade row designer is to tailor the blade shapes so that adverse gradients that produce large growths in the boundary layer are either reduced or controlled. It is of great concern to determine how well a particular analysis code predicts the surface velocities in order to determine how much faith the designer should place in the ability to tailor the blading for low losses. Blade loading in terms of differences in relative Mach numbers across the channel in the pitchwise direction versus percent chord are shown in figure 6 for 10%, 30% and 70% spans for both peak efficiency and near stall. The LA data is taken at the first point off the blade surface where a sufficient number of measurements were available. This was generally close to 5% and 95% pitch and, therefore, the analysis results were selected at these same locations. For all three spanwise locations the agreement on loading levels is reasonably good at peak efficiency. At near stall, however, the analysis yields much higher loadings at 10% and 30% span but is reasonable at 70%. The major difference in the loading level results from lower computed Mach numbers on the pressure surface than measured. Suction surface agreement is generally good for both flow rates. For all percent span locations and flow points, the shock location on the suction surface is well-predicted by the analysis.

CONCLUDING REMARKS

The comparisons between analytical and experimental data for a transonic, low-aspect ratio, axial-flow fan rotor have, in general, shown good agreement for all of the operating conditions presented. The overall trends in the relative Mach number contour maps indicate the very good shock capturing ability of the analysis code at all rotor span locations. The results presented have demonstrated the extent to which an inviscid Euler solver with a very simple boundary layer correction scheme was able to predict the complex 3-D flow field of this transonic rotor flow field. The code predicted the maximum flow for the rotor extremely well. The chordwise plots of relative Mach number reveal some areas in which analysis and LA data deviate. Some of these conditions may be traced to the boundary layer model used.

A three-dimensional analysis technique, such as the one used herein, would probably not be used in the preliminary design of turbomachinery

(a) 10% span from tip, peak efficiency

(b) 10% span from tip, near stall

(c) 30% span from tip, peak efficiency

(d) 30% span from tip, near stall

(e) 70% span from tip, peak efficiency

(f) 70% span from tip, near stall
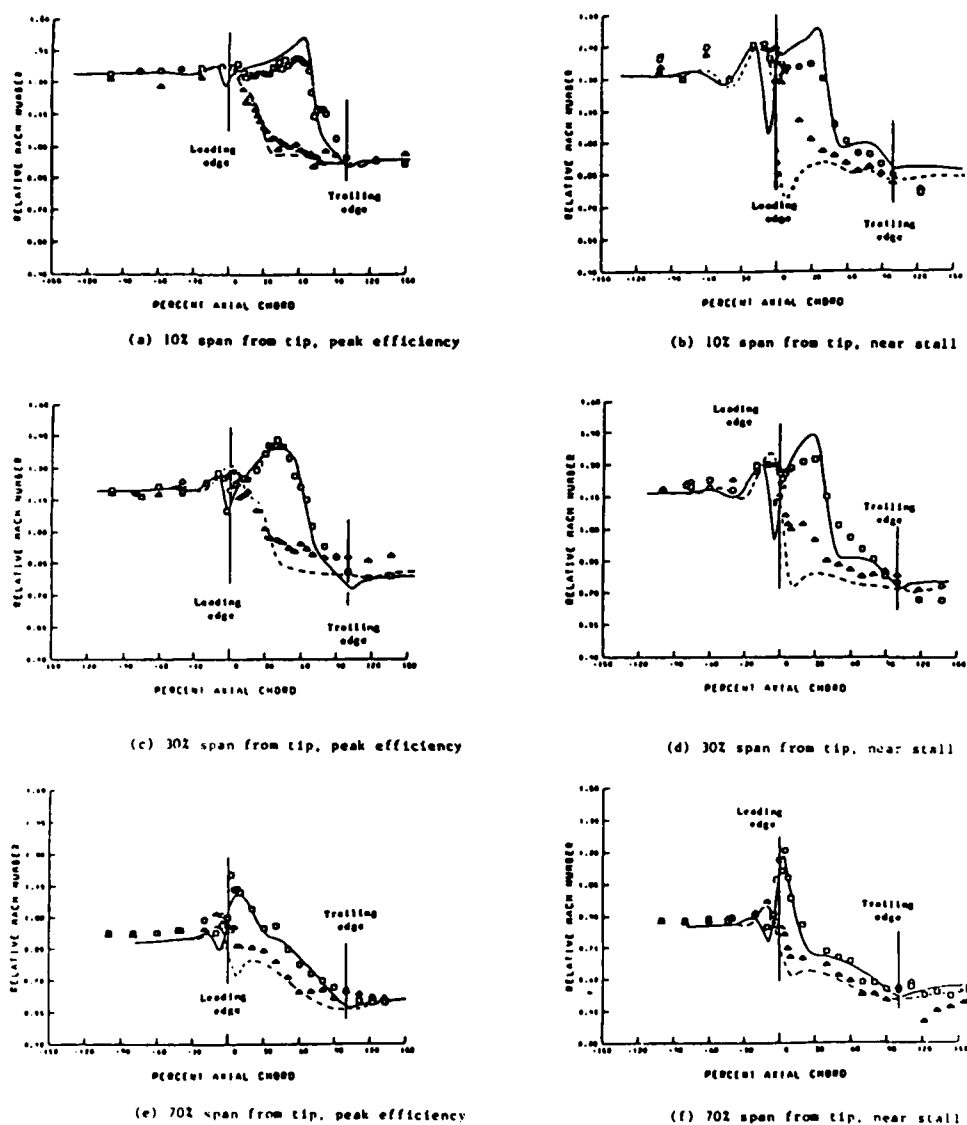
Figure 6. Variation of relative Mach number with percent chord for three spanwise locations.

☐ - LA data near suction surface
△ - LA data near pressure surface
—— Analysis results near suction surface
---- Analysis results near pressure surface

blading, however it may well be used to check a final configuration for design and off-design operation. This capability to compute the off-design flow field is extremely important and has shown to yield reasonably good quantitative results for the present rotor configuration. The analysis code did not predict the multiple shock system at 10% span, near stall on the pressure surface which was required to turn the flow parallel to the blade. This may have been due to the numerical smoothing used in the code and should be investigated further. Results concerning the intra-blade flow turning and work distribution should be useful information for compressor design systems.

As mentioned, the goal of the current NASA/ARMY research program is centered around increasing the fundamental understanding of the three-dimensional flow field of compressor blading. The material presented has been one such attempt to investigate the flow field of a particular transonic fan rotor and as such, the results shown apply only for this case. Further work dealing with different types of blading is needed to determine how much of what was observed experimentally and computed numerically for this isolated rotor, is generic to this class of turbomachine and how much was rotor specific. As more machines are investigated , questions such as blade row interactions and multistage matching should also be addressed.

REFERENCES

1. Chima R.V. and Strazisar A.J. , "Comparison of Two and Three Dimensional Flow Computations with Laser Anemometer Measurements in a Transonic Compressor Rotor", ASME Journal of Engineering for Power, Vol. 105, No. 3, July 1983, pp. 596-605.

2. Sarathy, K. P., "Computation of Three-Dimensional Flow Fields Through Rotating Blade Rows and Comparison with Experiment," ASME Journal of Engineering for Power, Vol. 104, No. 2, April 1982, pp. 394-402.

3. Singh, U. K., "A Computation and Comparison with Measurements of Transonic Flow in an Axial Compressor Stage with Shock and Boundary Layer Interaction", ASME Journal of Engineering for Power, Vol. 104, No. 2, April 1982, pp. 510-515.

4. Denton, J. D., Lecture Notes, ASME Turbomachinery Institute Course on Fluid Dynamics of Turbomachinery, July 18-27, 1983.

5. Pierzga, M.J. and Wood, J.R.,"Investigation of the 3-D Flow Field Within A Transonic Fan Rotor: Experiment and Analysis," ASME Paper No. 84-GT-200, 1984.

6. Strazisar, A. J., "Investigation of Flow Phenomena in a Transonic Fan Rotor Using Laser Anemometry," ASME Paper No. 84-GT-199, 1984.

7. Urasek, D.C., Gorrell, W.T and Cunnan, W.S., "Performance of Two-Stage Fan Having Low Aspect Ratio First-Stage Rotor Blading,"

NASA TP-1493 and AVRADCOM TR 78-49, August 1979.

8. Cunnan, W.S., Stevans, W. and Urasek, D.C., "Design and Performance of a 427-Meter-Per-Second-Tip-Speed Two-Stage Fan Having a 2.40 Pressure Ratio," NASA TP-1314, October 1978.

9. Powell, A.J., Strazisar, A.J. and Seasholtz, R. G., "Efficient Laser Anemometer For Intra-Rotor Flow Mapping in Turbomachinery," ASME Journal of Engineering for Power, Vol. 103, No. 2, April 1981, pp. 424-429.

10. Strazisar, A. J., and Powell, J. A., "Laser Anemometer Measurements in a Transonic Axial Flow Compressor Rotor," ASME Journal of Engineering for Power, vol. 103, No. 2, April 1981, pp. 430-437.

11. Denton, J. D., "An Improved Time Marching Method for Turbomachinery Calculation", ASME Paper 82-GT-239, 1932.

UNIVERSAL LITHIUM BATTERIES

EDWARD J. PLICHTA, Mr., AND MARK SALOMON, Dr.
US Army Electronics Development and Devices Laboratory
Power Sources Division
Fort Monmouth, NJ 07703

INTRODUCTION

The concept of a "Universal Lithium Battery" has emerged in response to the continuing proliferation of military batteries. In essence, the concept of Universal Batteries reduces to the identification of one or more chemistries which can be applied to a small number of battery configurations. The Universal Battery thus will serve both as a primary (i. e., throw-away) energy source for battlefield use, and as a rechargeable energy source for peacetime training purposes. It is also envisaged that a given Universal Battery can be used in a large number of configurations thereby reducing the proliferation of types of batteries.

In identifying one or more chemistries for Universal Batteries, lithium has been chosen as the anode material because of its very high energy density. Having selected a lithium anode, our major research has been directed towards the selection of suitable combinations of solvents, electrolytes, and cathodes. The solvent-electrolyte combination is an area of prime importance for numerous reasons: it must be highly conductive over a wide temperature range (-40°C to 70°C), and lithium must be stable in this system. The choice of cathode is also extremely important: it must be comparable to lithium in energy density, it must be able to undergo at least 25 discharge cycles, and it must be capable of delivering high currents required by many military devices.

One of the chemistries which appears to meet the above requirements involves the use of organic electrolytes with an intercalating cathode such as TiS_2. The rechargeable cell thus can be described as

$$Li/organic\ electrolyte/TiS_2 \qquad [1]$$
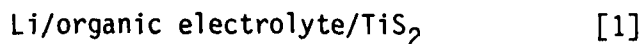
In cell [1] the anode reaction is

$$Li \rightleftarrows Li^+ + e^-$$

29
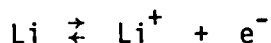
and the cathode reaction (involving intercalation of lithium) is

$$TiS_2 + Li^+ + e^- \rightleftharpoons LiTiS_2$$
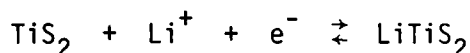
The overall cell reaction is thus
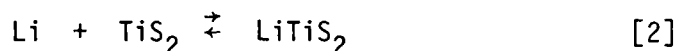
$$Li + TiS_2 \rightleftharpoons LiTiS_2 \qquad [2]$$

It is also important to note that because Li_xTiS_2 is a true solid solution over the composition range of $0 \leq x \leq 1$, its chemical potential is a function of lithium content, and hence the open circuit cell potential is a function of the state-of-charge: i.e., a simple measurement of the cell potential will provide exact values of the remaining cell capacity.

SOLVENT STUDIES

The requirements of high power densities over a wide temperature range translates into the need to develop new electrolytes. Organic solvents (required for lithium stability) typically have low conductivities, and electrolytes are generally characterized by low solubilities. However, in recent publications (1-3) we have reported that the use of mixed organic solvents and $LiAsF_6$ electrolytes results in electrolyte solutions in which the electrolytic conductivities are much greater than in the pure organic solvents. In Figure 1 we show our recent results for the electrolytic conductivities (κ in units of S cm^{-1}) for 1 mol dm^{-3} solutions of $LiAsF_6$ and $LiAlCl_4$ mixtures of 4-butyrolactone (4-BL) with dimethoxyethane (DME) and tetrahydrofuran (THF). At 25°C the maximum in the electrolytic conductivity for these systems is around 0.018 S cm^{-1} which occurs in a solvent containing between 65-75 mass % ether (DME or THF). The use of highly concentrated solutions is required to provide high conductivities, but previous studies (e.g., refs. 4, 5) using 1.5 - 2.5 mol dm^{-3} $LiAsF_6$ in pure ethers such as 2-methyl tetrahydrofuran (2-Me-THF), pure THF, or mixtures of 2Me-THF with THF resulted in very poor low temperature behavior because of the rapid decrease in the solubility of $LiAsF_6$. However, the use of 4-BL which has a moderately high dielectric constant of 41 D (2) promotes increased solubilities in mixtures with THF or DME, and allows us to use much lower $LiAsF_6$ or $LiAlCl_4$ concentrations. The use of lower electrolyte concentrations thus prevents the precipitation of the salt as the temperature is decreased. At the same time, the use of 4-BL mixtures with DME or THF leads to electrolytic conductivities much higher than for pure ethers or mixed ether solutions. For example, Table 1 shows our recent results for the electrolytic conductivity of 0.8 mol dm^{-3} $LiAsF_6$ in DME containing 24.4 mass % 4-BL. The results in Table 1 are significantly superior to previous results for 1.5 mol dm^{-3} $LiAsF_6$ in a 50% mixture of 2Me-THF/THF in which sufficiently conductive solutions could not be realized below -20°C (5).

On the basis of the results for conductivities in mixed organic sol-
vents, we have selected two mixtures for use in small laboratory recharge-
able cells. The mixtures are: 24.4 mass % 4-BL in DME, and 35 mass % 4-BL
in THF. The experimental results given below correspond to cells using the
4-BL/DME mixed solvent and other mixtures presently being investigated will
be reported at a future date.

We have also addressed the question of why these mixed solvents have
larger conductivities than either of the pure components. The so-called
"synergistic" effect is a result of numerous factors involving viscosity-
dielectric constant effects, but the major factors appear to be due to the
unique ability of ethers to strongly coordinate small cations such as Li^+
(1-3). The conductivities of electrolytes (MX) in organic solvents are
generally very low due to low solubilities and the formation of "strong"
ion pairs in solution according to

$$M^+ + X^- \rightleftarrows (M^+X^-) \qquad [3]$$
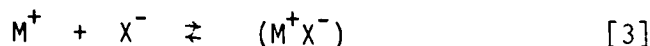
The formation of ion pairs (M^+X^-) in effect removes charge carriers from
solution. By using ethereal solvents, Li^+ is so strongly coordinated (e.g.
see refs. 2-3) that the equilibrium in e.g. [3] is shifted to the left, and
lithium salts become highly soluble.

CELL STUDIES

Having selected the solvent mixture of 24.4 mass % 4-BL in DME, the
cathodes were selected on the basis of previous studies (4-5). TiS_2 (see
eqns. [1] and [2]) was selected because of its ability to undergo hundreds
of charge-discharge cycles, and MoS_3 was selected because of its high the-
oretical energy density. TiS_2 can intercalate a maximum of one equivalent
of Li, but MoS_3 can intercalate a maximum of four equivalents of Li shown
by the cell reaction (where x = 4)

$$xLi + MoS_3 \rightleftarrows Li_xMoS_3 \qquad [4]$$
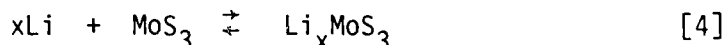
However, experimentally (6) it is found that the maximum value of x = 4 is
approached $(x \cong 3.5)$ at room temperature at very low current densities
$(\leq 0.5$ mA cm$^{-2})$. Military equipment generally require cells to be capable
of delivering 2-5 mA cm^{-2}, but at these current densities the maximum value
of x in Li_xMoS_3 drops to around 1-2 equivalents of Li. At low temperatures
$(\leq -20°C)$ the capacity of the Li/MoS_3 cell drops to zero. The major cause
of this rapid fall-off in capacity is attributed to high cell resistance
due to the fact that MoS_3 is a non-conductor as opposed to TiS_2 which is a
semi-conductor. Thus, to increase the capacity of MoS_3 cathodes it is
necessary to increase its ability to sustain large current drains, and one

method we explored was to incorporate TiS_2 in MoS_3. A mixed TiS_2-MoS_3 cathode containing 20 mass % TiS_2 was found to provide high current densities at ambient temperatures, but at low temperatures (-30°C) the capacity of the Li/TiS_2-MoS_3 dropped to that of Li/TiS_2. In fabricating the cathodes, the following mass % compositions were employed:

TiS_2 cathodes: 83 % TiS_2, 85 % teflon, 8.5 % carbon

MoS_3 cathodes: 70 % MoS_3, 12 % teflon, 18 % carbon

Mixed cathodes: 60 % MoS_3, 20 % TiS_2, 10 % teflon, 10 % carbon

The experimental cell discharge/charge results are summarized in Figures 2-5. Figure 2 shows initial 2 mA cm^{-2} discharge curves for lithium cells at 25°C comparing TiS_2, MoS_3-TiS_2, and MoS_3 cathodes: the maximum number of moles of lithium intercalated in each of these cathodes are 0.86, 2.30 and 2.99, respectively. Increasing the current density to 5 mA cm^2 results in the behavior shown in Figure 3, and for the cathodes TiS_2, MoS_3-TiS_2, and MoS_3 the maximum number of lithium equivalents which are intercalated drops to 0.81, 2.00, and 1.95, respectively. At this higher rate (at 25°C) the mixed cathode MoS_3-TiS_2 performs better than pure MoS_3 due mainly to its higher operating voltage which results from the incorporation of the semi-conducting material TiS_2 into the cathode. In spite of the high resistivity of pure MoS_3, this cathode can still operate at 5 mA cm^{-2} at ambient temperatures when using highly conductive electrolyte solutions. At low temperatures, the capacity of pure MoS_3 drops so rapidly that we have not pursued its characteristics. However, the mixed MoS_3-TiS_2 cathode appears to operate well down to -30°C at 2 mA cm^{-2} as shown in Figure 4. The number of equivalents of lithium intercalated at -10°C, -20°C and -30°C are 1.70, 1.13, and 0.81, respectively. At this moderate current density of 2 mA cm^{-2}, the mixed MoS_3-TiS_2 cathode behaves similar to that of pure TiS_2. The discharge characteristics of the Li/TiS_2 cell at low temperatures are shown in Figure 5. In this figure the number of lithium equivalents intercalated per mole of TiS_2 are 0.77, 0.66, 0.53, 0.33, 0.12, at 0°C, -10°C, -20°C, -30°C, and -40°C, respectively.

At 5 mA cm^{-2} at -30°C both Li/TiS_2 and Li/MoS_2-TiS_2 cells have negligible capacity: in both cases about 0.07 equivalents of lithium are intercalated.

CONCLUSIONS

A number of new mixed solvents have been shown to improve the current capabilities and low temperature behavior of secondary lithium cells. At very low temperatures and very high current densities, cell capacities drop significantly. This behavior is due to a combination of moderate solvent resistance and the high resistance of the cathode materials. Future work

to develop rechargeable lithium cells will be directed towards improving the conductivity of intercalating cathodes, and towards the identification of new solvent mixtures.

REFERENCES

1. M. Salomon and E. J. Plichta, Paper Presented at the Spring Meeting of the Electrochemical Society, Extended Abstract No. 807, San Francisco, CA, 1983.

2. M. Salomon and E. J. Plichta, Electrochimica Acta, 28, 1681 (1983).

3. M. Salomon and E. J. Plichta, Electrochimica Acta, accepted for publication.

4. K. M. Abraham, J. Power Sources, 7, 1 (1981/82).

5. K. M. Abraham and J. L. Goldman, J. Power Sources, 9, 239 (1983).

6. A. J. Jacobson, R. R. Chianelli, S. M. Rich, and M. S. Whittingham, Mat. Res. Bull., 14, 1437 (1979).
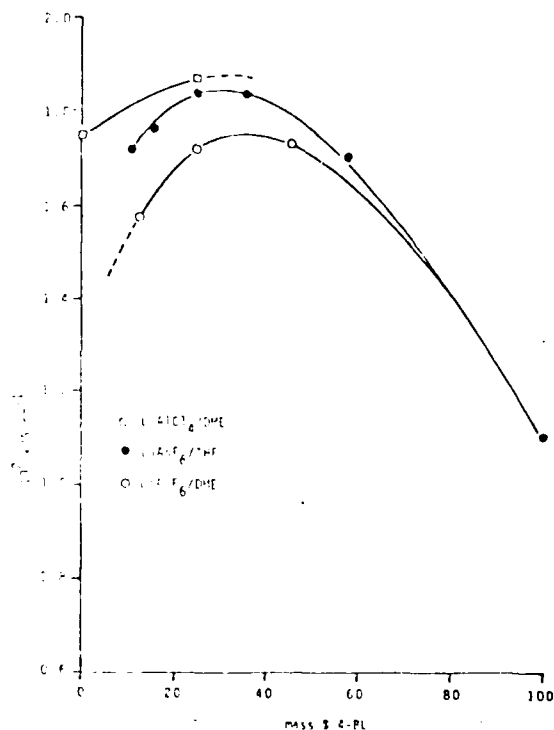
Fig. 1. Electrolyte conductivities of 1M solutions in 4-BL mixtures with DME and THF.
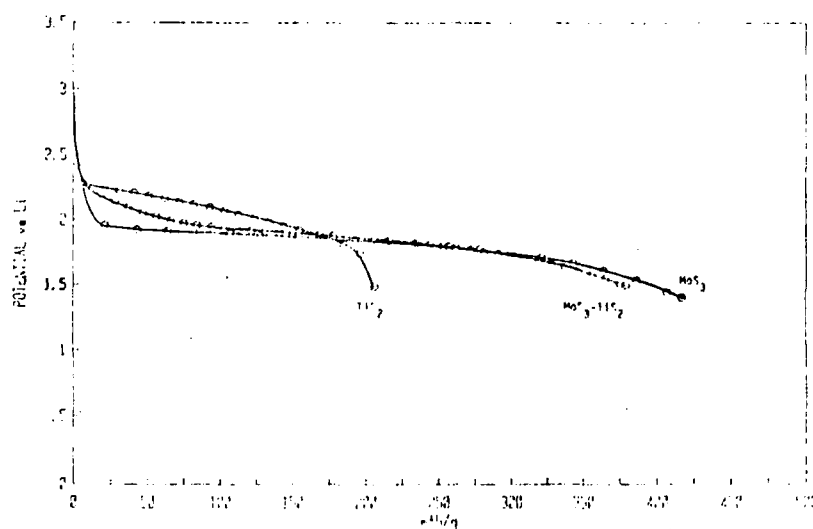
Fig. 2. Discharge at 2 mA cm^2 at 25°C. Electrolyte is 1M LiAsF$_6$.
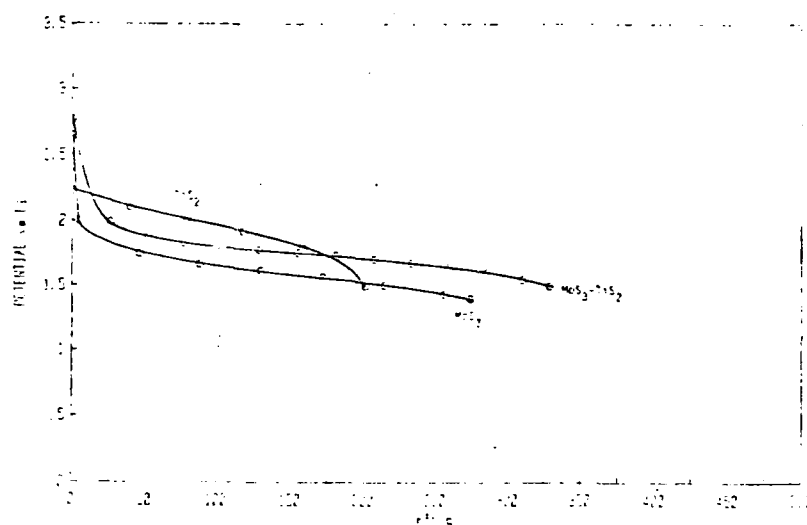
Fig. 3. Discharge at 5 mA cm^{-2} at 25°C. Electrolyte is 1M LiAsF$_6$.
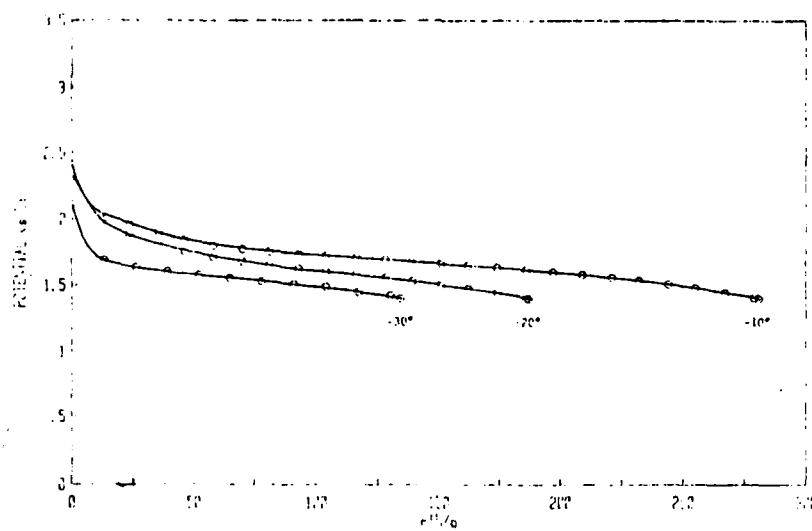
Fig. 4. Discharge of Li/MoS$_3$ - TiS$_2$ cell at 2 mA cm^{-2} in 0.8M LiAlCl$_4$.
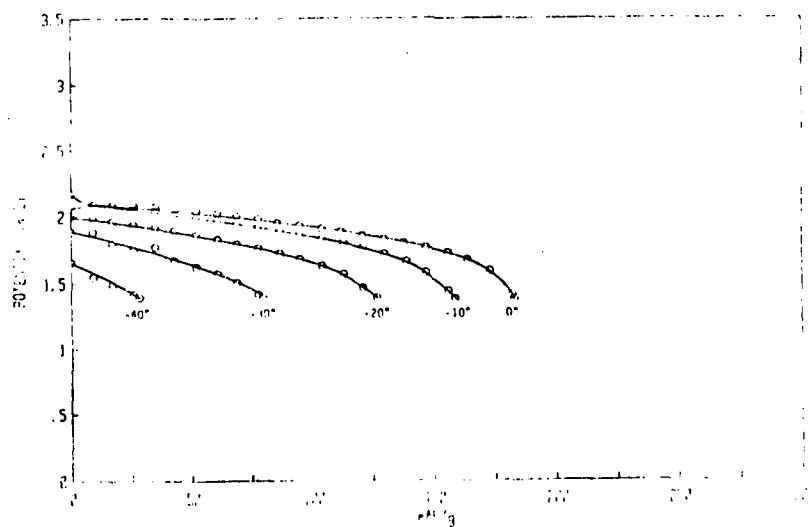
Fig. 5. Discharge of Li/TiS$_2$ cell at 2 mA cm^{-2} in 0.8M LiAlCl$_4$.

Table 1. Electrolyte Conductivity of 0.8 mol dm^{-3} LiAsF$_6$ in 24.4 % 4-BL in DME as a Function of Temperature.

t/°C	κ/S cm^{-1}
25	0.018
-10	0.0097
-20	0.0070
-30	0.0044

A Momentum Analyzer for Intense
Relativistic Electron Beams (U)

*Maj. Edward W. Pogue
Los Alamos National Laboratory
Los Alamos, NM 87545

SUMMARY

A momentum analyzer has been built and tested on the PHERMEX standing-wave rf electron linac. The analyzer has measured the 28-MeV 400-A PHERMEX beam to a resolution of 1.0 MeV and 350 ps. Comparison of the measured beam energy structure with a computer model of the PHERMEX beam shows close agreement.

INTRODUCTION

PHERMEX is a three-cavity rf electron linac used to generate a short x-ray burst for flash x-radiographic measurements of exploding devices. The acronym PHERMEX stands for Pulsed High Energy Radiographic Machine Emitting X Rays. It consists of a hot-cathode injector powered by a 200-ns pulser, followed by three accelerating cavities where the beam is accelerated and chopped into a series of micropulses of 4-ns duration (Fig. 1). The electron beam is then transported through a series of solenoidal focusing lenses to a tungsten target (Fig. 2).[1] The cavities operate in the TM_{010} mode at 50 MHz and provide a peak electric field on axis of approximately 5 MeV/m. The electron beam is focused to a 3-mm spot on the target and converted to a high-quality x-ray source through the bremsstrahlung process. The small spot size and short duration of the PHERMEX beam allow resolution of very small density variations in very thick high-atomic-number exploding objects.

In recent years the increased interest in electron-beam propagation has prompted a series of experiments in that area on PHERMEX. The most significant result of these experiments is that no evidence of hose instability has been found.

The expected momentum distribution was first calculated during the design of the PHERMEX accelerator using the GRAPE SEED III computer code.[2] The results are shown in Fig. 3. To experimentally verify this momentum distribution, a bending-magnet-type momentum analyzer has been constructed for PHERMEX.
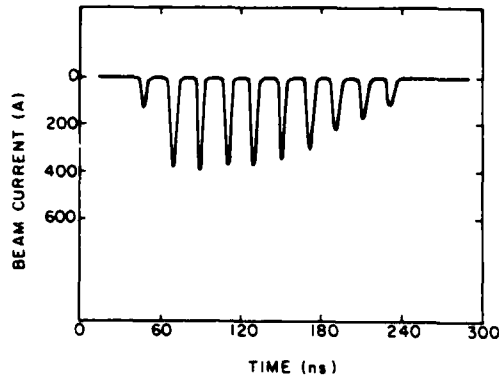
Fig. 1.   Beam current pulse train.
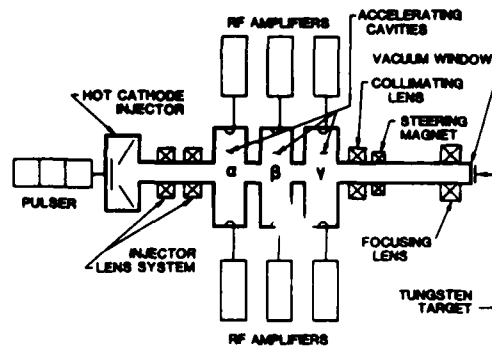
Fig. 2.   PHERMEX accelerator.
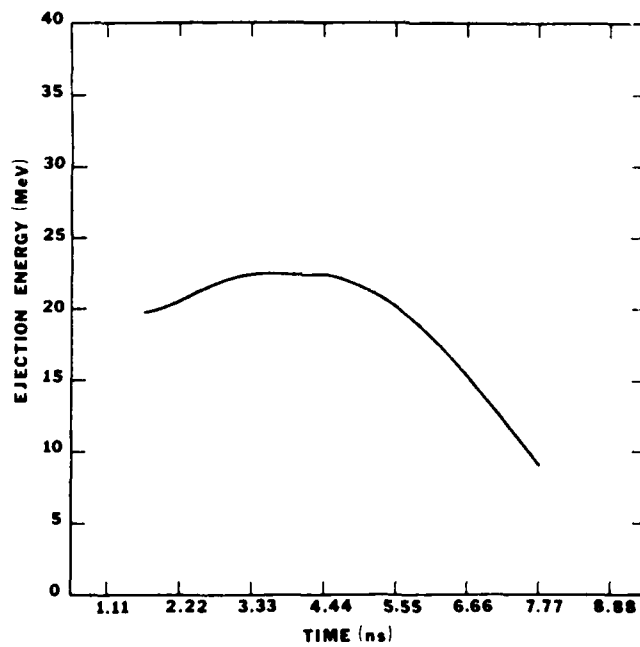
Fig. 3.   Micropulse energy variation.

THEORY

The theory of operation of this diagnostic is that when electrons of different momentums go through a magnetic field perpendicular to their line of flight, they are deflected by an angle inversely proportional to their momentum. Momentum (P) is defined for relativistic electrons by $P = (v/c)E$ where v is the electron velocity, c is the speed of light, and E is the total relativistic energy of the electron. The force of the magnetic field acting on the electron is $\vec{F} = q\vec{v}\times\vec{B}$ where q is the electron charge, $\vec{v}$ is the electron velocity, and $\vec{B}$ is the external magnetic field. Because all the electrons in the PHERMEX beam are highly relativistic, one can make the approximation that $v = c$, so that the force is the same for all electrons. Although all electrons feel the same force, their displacement under the influence of that force is proportional to their inertia or mass. Electrons at different energies have different masses as related by Einstein's theory of special relativity. (Mass = $m_0\gamma$ where m_0 is the rest mass and γ is the relativistic factor.) To determine the angular deflection of a given electron, the following equation[3] is used

$$\Delta\theta = \frac{\int\vec{B}\cdot d\vec{l}}{B_\rho}$$ (1)

where $\Delta\theta$ is the angular deviation in radians, $\vec{B}$ is the external applied magnetic field, $d\vec{l}$ is an increment of path length, and B_ρ is the magnetic rigidity of the particle related to the momentum by $B_\rho = P(MeV)/0.3$ and is in units of kG-cm. Thus for a higher momentum, $\Delta\theta$ is less.

By directing the beam through the analyzing magnetic field of this instrument and by measuring the deflected beam current at a particular angle ($\Delta\theta$) as a function of magnetic field strength, the time-resolved beam momentum distribution can be determined.

The beam peak energy can be computed from the following equation:

$$E(MeV) = (2.6 \times 0.72)(E_\alpha + E_\beta + E_\gamma) + V_g$$ (2)

where E is the measured beam energy in MeV, E_α is the measured field strength in alpha cavity in MeV/m, 2.6 m is the cavity length, 0.72 is the cavity-acceleration efficiency, and V_g is the injected electron energy in MeV.

39

In the following sections, the apparatus design, experimental procedure, data analysis, sensitivity, and results are given.

APPARATUS DESIGN

The analyzer consists of collectors to measure the deflected electron current, an electromagnet to provide an adjustable magnetic field, and a vacuum chamber to collimate and provide a low-scattering environment for the electron beam.

The collector was designed to cover the 20- to 60-MeV momentum range of interest. The collector must have a frequency response greater than 1000 MHz, attenuate the reflected signals, and be sensitive to currents of 1 A. The collector must also be able to accomplish this in a high-rf-noise environment.

The unique collector design chosen incorporates a rectangular center conductor that is sufficiently thick to stop all the electrons; around this is a dielectric spacer and then a conducting box that completely encloses the collector and grounds to the vacuum chamber. The conducting box shields the collector from the rf noise and creates a square coaxial line that has a characteristic line impedance. One end of the collector is terminated into a 50-ohm connector, whereas the other end is grounded to the conducting box through a resistor array that matches the line impedance. To keep the collector small, a material with good stopping power (lead) was chosen. Kapton was picked as the dielectric because of its durability under electron bombardment. The collector is shown in Fig. 4. At 2.54-cm thick, the lead collector is two ranges for a 50-MeV beam. A 3.3-ohm resistor array was chosen to match the line impedance of the collector. For a square geometry this is given by

$$Z = \frac{94.15}{(C/\epsilon + w/t)} \tag{3}$$

where C is the capacitance to ground, ϵ is the Kapton dielectric constant (3.4), w is the width of the collector, and t is the thickness of the Kapton. By matching impedances, the reflected signal is kept to a minimum. A copper box encloses all but one side of the collector and is mounted directly to the vacuum chamber to insure a good return path for the electrons.

In operation, a narrow beam (5-mm diam) of electrons passes through the 1 mm of Kapton and strikes the center of the lead collector where it is absorbed (Fig. 4). This generates a signal that propagates to both ends of the collector. At the resistor-array end of the collector, the matching impedance of the line and resistor array insure that the signal is absorbed

ELECTRON BEAM

LEAD COLLECTOR

TYPE N
CONNECTOR

3.3 ohm RESISTOR ARRAY

THIS SECTION WRAPPED
IN KAPTON , 1mm THICK.
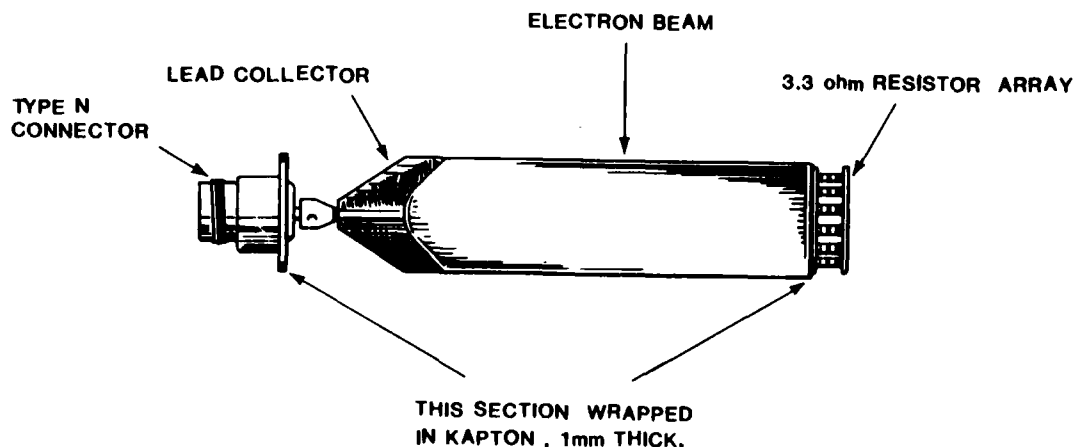
Fig. 4. Collector assembly.

with little or no reflection. The signal in the opposite direction pro-
vides the signal that is measured thr(ɪgh the 50-ohm Type-N connector.
Because of the impedance mismatch, the majority of this signal is reflected
back to the 3.3-ohm array and absorbed, thus preventing ringing of the
collector.

A Varian 4005 magnet was used. The magnet has 10.16-cm flat
cylindrical pole faces. Both the power supply and the magnet are water
cooled. The pole faces and windings are shown in Fig. 5. A precision
current shunt was installed, and a digital voltmeter was used to measure
currents to ±0.05 A. Because of the width of the windings, the fringe
fields remain significant out to a much greater radius than the pole face.
The 31.8-mm gap between the poles was chosen, based on the width of the
interior of an available vacuum chamber. When the magnetic-field profiles
are normalized to peak field at each current setting, the profile shape
remains constant to within ±2%. Only above 10 kG are the pole pieces
starting to saturate.

Because of the large fringe field outside the pole-face area, the path
of the electrons could not be easily approximated. To deal with this prob-
lem, a computer program (MAGNET³) was written by L. A. Wright of Mission
Research Corporation to determine the path of electrons as they pass
through the field of the bending magnet. Using Eq. (1), an electron of a
given momentum value is stepped through the magnetic fields 1 mm at a time,
with its new position and new direction computed at each step. The results
of Wright's calculations for a collector at 45° can be seen in Fig. 6. The
results are for a uniform 1-mm beam spot size at the focus 20 cm from the
center of the pole face, with 18-mrad divergence, and indicate that a reso-
lution of 0.6 MeV should be achieved.
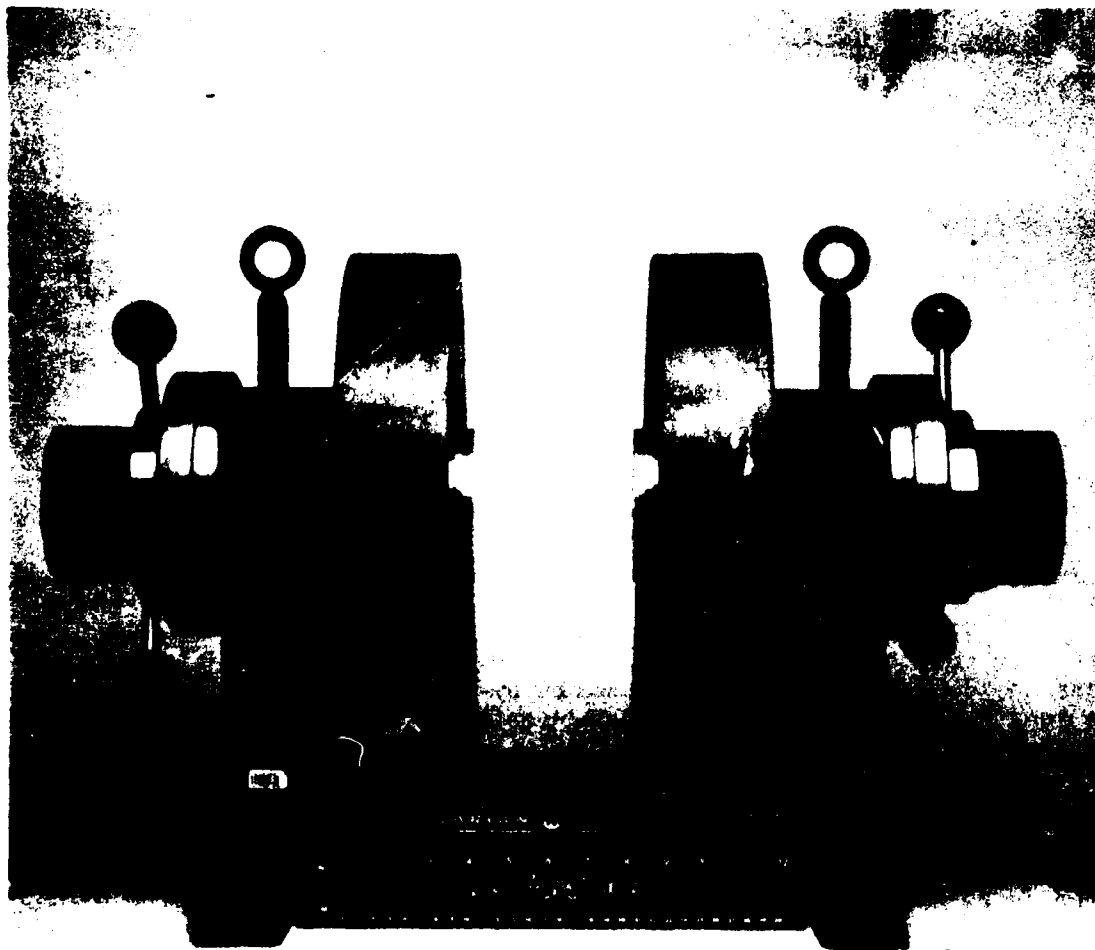
Fig. 5. Varian 4005 magnet.
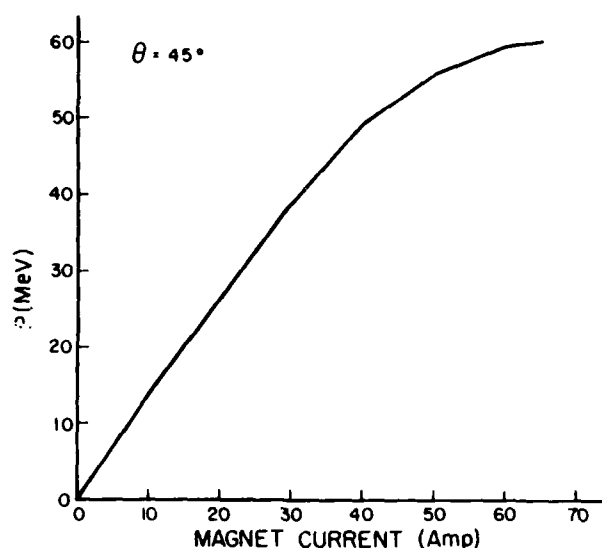
Fig. 6.   Result of MAGNET computer code.

An existing vacuum chamber was modified for use with the new collectors
and to facilitate its use with PHERMEX. Collectors were placed at 0 and
45° to the beam axis (Fig. 7). A 1-in.-thick lead collimator with a
5-mm-diam hole was used to define the electron beam entering the collector.
By assuming a 1-mm spot size and a 40-MeV peak energy beam, the size of the
hole was determined by the use of the TRANSPORT computer code.[*]

To insure that anomalous electrons, those with large divergences, do
not contribute to false signals, collimation was added in the form of a
circular aluminum plate. Based on the computer-generated path of the elec-
trons that would be accepted at the 45° collector, a channel was cut in the
aluminum plate.

To easily mate the momentum analyzer to the PHERMEX drift tube, a Die
table was used. The hydraulic height adjustment and the locking wheels of
the Die table allowed rapid emplacement and alignment of the momentum
analyzer at the firing point.

To couple the vacuum chamber to the PHERMEX drift tube, a copper
adapter sleeve was designed. The design incorporated a self-integrated
B-dot loop for an accurate measure of the beam current entering the
momentum analyzer. This sleeve also insured a good return-current path.
The beam traverses 4.5 cm of air while passing through the sleeve.
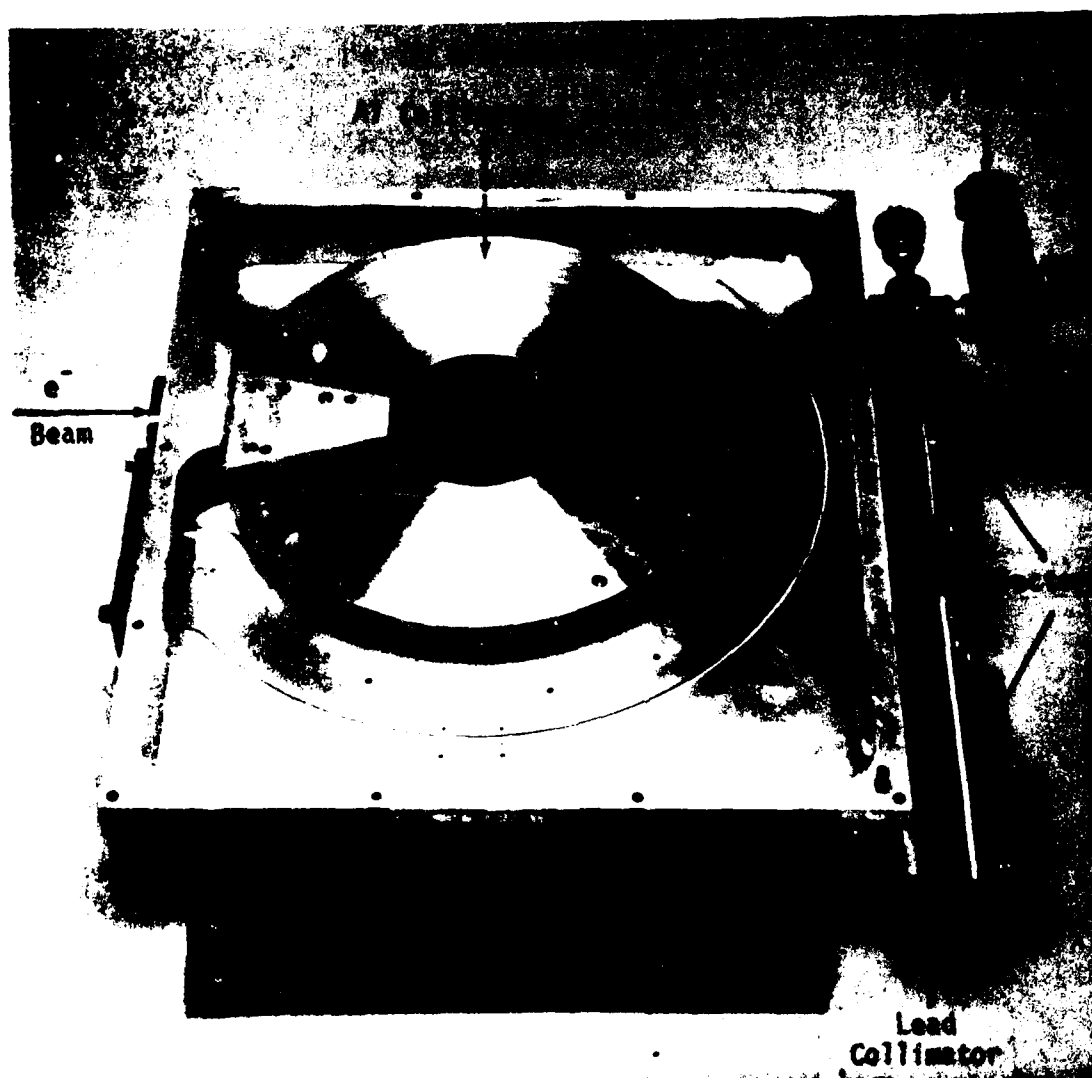
Fig 7. Vacuum chamber with collectors mounted.

EXPERIMENTAL PROCEDURE

A Physics International Faraday cup in a short drift section was initially used to tune the accelerator. Glass witness plates were taken at the exit point of the beam and at the Faraday cup. These plates were used to establish the beam path and orient an optical alignment scope. Examination of the glass plates taken at the exit foil reveals that a 1- to 2-mm spot size was achieved. The next step was to install and align the momentum analyzer on the end of the PHERMEX drift tube.

The signals from the B-dot loop and the collector were recorded on Tektronics Model 7104 scopes and on Tektronics R-7912 transient recorders. This provided two records of each signal, giving the ability to look at both the macropulse and one of the micropulses during the same pulse.

The data-collection effort fell into two catagories: measurements of the momentum distribution of the micropulse, and a check of the sensitivity of the measurements.

In the first measurement, the peak momentum was established, and the magnet current was then stepped down in 0.4-A steps until the minimum momentum was found. The 0.4-A steps corresponded to 0.6-MeV changes in the momentum of the electrons. Data were recorded at 1 ns per division on the third micropulse in the train.

The sensitivity of the data was checked by decreasing the final focusing magnet's field while using the straight-through collector and by applying small current variations to the bending magnet to scan the radial distribution and check for x-ray effects. Sensitivity to the accelerator steering lenses was also checked.

DATA ANALYSIS

The analysis of the data was conducted on both the micropulse and macropulse information. The data were digitized using an HP9831A computer and stored for use with integration routines and various other analytical programs.

The peak momentum measured was 28.6 MeV. In Fig. 8, a representative sample of the raw data taken on the third micropulse is given. A careful examination of the expected distribution shown in Fig. 3 shows that the results were, in general, as predicted.

To check the rate of cavity energy depletion, the peak momentums for the third and eighth micropulses were compared. For the third, the value was 28.6 MeV; and for the eighth, the value was 26.7 MeV. This gave a cavity depletion rate of 1.4% per micropulse.

The last part of the analysis consisted of digitizing micropulse data and performing computer manipulation of the data. Included in this data was a complete sweep of the beam going from peak-to-minimum momentum in 0.5-MeV steps. Through the use of the HP9831A computer, the matrix of momentum, time, and current for the third micropulse was stored. Taking
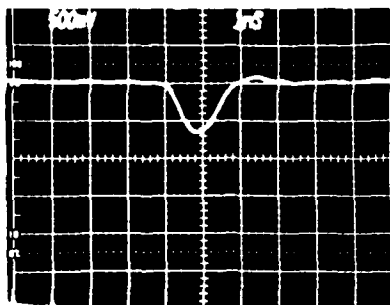
28.1 MeV
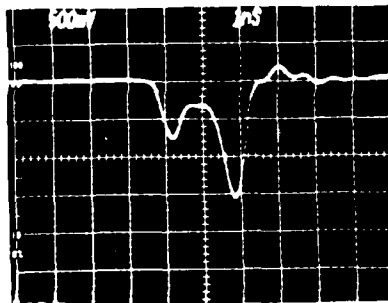
26.4 MeV
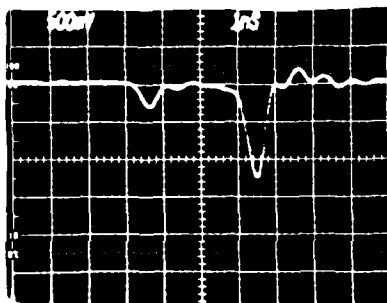
24.8 MeV
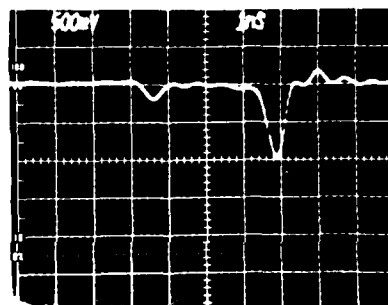
23.7 MeV
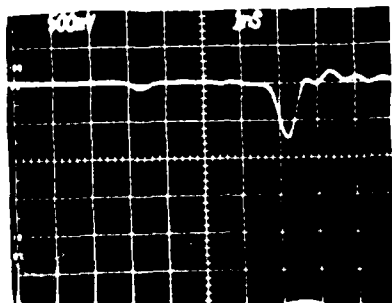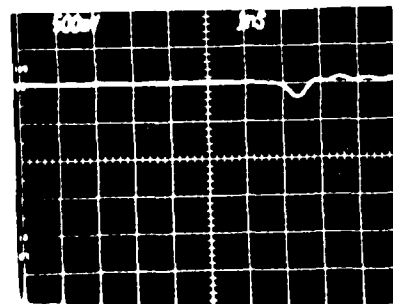
1.1 MeV

pulse.

this three-dimensional parameter space and plotting the data in two dimensions gives: Fig. 9, Charge-vs-Momentum; Fig. 10, Momentum-vs-Time; and Fig. 11, Charge-vs-Time. In the Momentum-vs-Time plot, it is interesting to note that the measurement disagrees with D. Mitrovich's computer model of the accelerator (PHERMEX[5]) at the head and tail of the micropulse. The theory line represents the output of the third cavity and does not take into account the losses due to steering and focusing through the 10-m length of drift tube from the exit of gamma cavity to the entrance of the momentum analyzer. In Fig. 10, Charge-vs-Time, the reconstructed line was computed by summing the 17 individual pulse forms from the momentum analyzer to attempt to reconstruct the current pulse form given by the B-dot and Faraday-cup signal. Overall, the agreement is adequate.

The error analysis for the momentum analyzer was done for alignment, displacement, and acceptance through the use of the computer code MAGNET.[3] These factors combine to give an error of $\pm 2.46\%$ at the peak momentum. The only other error to be included is that of the gaussmeter used to map the momentum-analyzer bending magnet. This error was $\pm 1.0\%$, giving a total error of $\pm 3.46\%$, or for peak momentum, 28.6 $\pm$ 1.0 MeV.

In regarding the sensitivity measurements, none of the changes made a detectable change in the B-dot loop signal or the straight-through Faraday-cup signal. This shows that the measurements were relatively insensitive to minor variations in these settings.

CONCLUSIONS

Peak momentum of the PHERMEX electron beam was determined to be 28.6 MeV.

The momentum analyzer performed successfully and can be used with confidence to measure the PHERMEX beam at higher energies.

The results of these measurements are being used to update computer codes that model beam propagation.

ACKNOWLEDGEMENTS
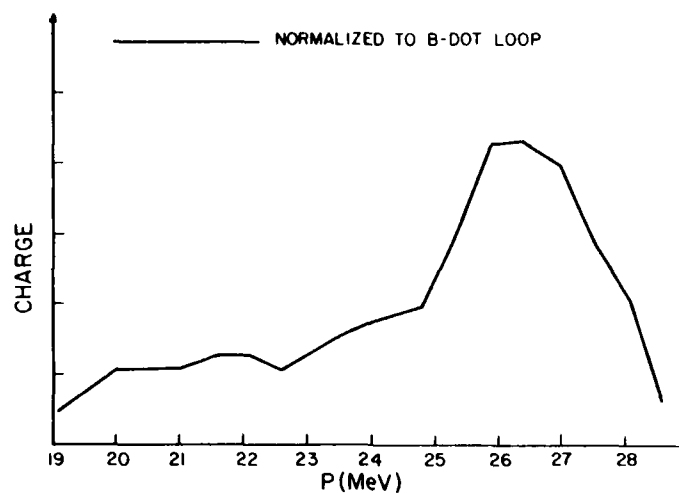
Fig. 9.   Charge-vs-momentum.
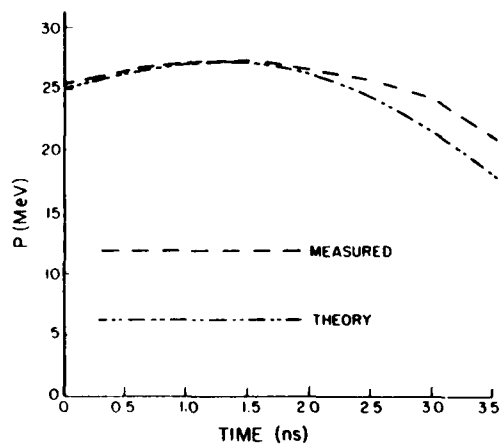
Fig. 10.   Time-vs-momentum.
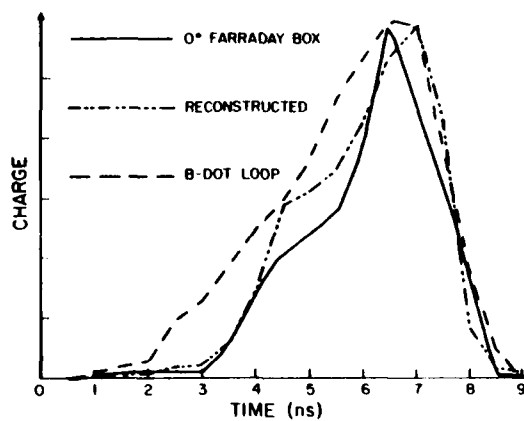
Fig. 11.   Charge-vs-time.

REFERENCES

1. T. P. Starke, "PHERMEX Standing-Wave Linear Electron Accelerator," Los Alamos National Laboratory, IEEE Transaction on Nuclear Science, Vol. NS-30, No. 2, (April 1983).

2. D. Venable, D. O. Dickman, J. N. Hardwick, E. D. Bush, Jr., R. W. Taylor, T. J. Boyd, J. R. Ruhe, E. J. Schneider, B. T. Rogers, and H. G. Worstell, "PHERMEX: A Pulsed High-Energy Radiographic Machine Emitting X-Rays," Los Alamos Scientific Laboratory report LA-3241.

3. L. A. Wright "MAGNET: A Simple Code to Calculate Electron Trajectories in a Spatially Varying Magnetic Field," Mission Research Corporation report AMRC-N-237 (September 1983).

4. K. L. Brown, F. Rothacker, D. C. Carey, and Ch. Iselin, "TRANSPORT: A Computer Program for Designing Charged Particle Beam Transport Systems," Stanford Linear Accelerator Center report SLAC-91 (May 1977).

5. D. Mitrovich, "Calculation of Electron Trajectories Through the PHERMEX Cavities," Mission Research Corporation report AMRC-N-240 (September 1983).

NONLINEAR OPTICAL PROPERTIES OF SELECTED ORGANIC MATERIALS (U)

JUERGEN L. W. POHLMANN, DR., *EDWARD J. SHARP, DR.
AND GARY L. WOOD MR.
NIGHT VISION AND ELECTRO-OPTICS LABORATORY
FORT BELVOIR, VIRGINIA 22060

INTRODUCTION

The study of nonlinear optics has grown enormously since the 1960's and many uses have already been devised for exploiting nonlinear optical effects both for the study of material systems and for the development of coherent light sources. Note that this implies that there are two distinct ways to look at the nonlinear interaction of electromagnetic radiation with matter. In one case the effect of the interaction on the radiation field itself is of primary interest, such as providing new coherent radiation sources. On the other hand the effect of the interaction on the radiation field can be detected in most experimental arrangements in such a way that it can provide information on the reaction of the molecular system, for example, an ensemble of molecules can react to the radiation field in such a way that the detectable changes in the field can provide specific information regarding the molecular states and their dynamics. We are primarily concerned here with the latter of these two viewpoints and even more explicitly with the applications that stem from such interactions in organic compounds.

It is to be expected that most attention to date has been directed toward the higher order optical processes in semiconductors and inorganic dielectric insulators since there exists already a great wealth of information on these materials from earlier studies of optical, electrical, and thermal transport phenomena. This being the case, almost all of the commonly used nonlinear optical materials are inorganic solids.

The nonlinear optical properties of organic compounds are of interest both from a fundamental and a practical point of view. Recently this interest has become more evident as researchers turn their attention to the nearly unlimited number of organic compounds which are readily available for nonlinear optical applications. Nonlinear optical spectroscopy of molecular

51

systems is providing insight into the dynamics of molecular states in such fundamental processes as ionization and energy transfer (1). Large cptical nonlinearities have been observed in organic and polymeric crystals and films (2), (3), as well as in liquid crystals (4), (5).

NONLINEAR PHENOMENA

Generally, the realm of nonlinear optics is understood to include all the phenomena for which electromagnetic field intensities of higher powers than the first play a dominant role. The basis of this stems from the fact that for all paractical purposes the mathematical description of the majority of nonlinear optical processes can be described in a semiclassical way by using an electric-dipole approximation to the Hamiltonian. This requires that the polarizable unit of the material occupies a volume in which the electromagnetic field is uniform for any given time, i.e., the occupied volume is small compared to the radiation wavelength. In this situation the total polarization set up in the medium is a summation over all the polarizable units, and since the field is uniform within this unit the interaction looks like that of an electric-dipole in a constant field. Since we can use the electric-dipole approximation we can therefore describe much of nonlinear optics in terms of susceptibilities; or in other words, the induced optical polarization in the material can be represented as a series expansion in terms which are directly proportional to some power of the field amplitude.

$$P = \varepsilon_o \chi E + \varepsilon_o \chi^{(1)} E^2 + \varepsilon_o \chi^{(2)} E^3 + \varepsilon_o \chi^{(3)} E^4 + \cdots \qquad [1]$$

Here the relationship between the polarization and the applied field is most generally of tensor form and the susceptibilities are functions of the radiation frequencies and material parameters but are independent of position in the electric-dipole approximation.

The well known optical phenomena of reflection and refraction arise from the first term of the optical polarization, which is linear in the field amplitude. A wide variety (6) of interesting and important optical phenomena arise from the nonliner susceptibilities which are quadratic and cubic with respect to the field amplitude. For our purposes we can write Eq. [1] as a polarization $\vec{P}$ consisting of a linear $\varepsilon_o \chi \vec{E}$ and a nonlinear term $\vec{P}_{NL}$ as:

$$\vec{P} = \epsilon_o \chi \vec{E} + \vec{P}_{NL} \qquad [2]$$

where for an isotropic medium, since $\chi^{(1)}$ is zero we can write

$$|\vec{P}_{NL}| = \epsilon_o \chi^{(2)} \langle \vec{E} \cdot \vec{E} \rangle \vec{E} \qquad [3]$$

Here $\langle \vec{E} \cdot \vec{E} \rangle$ is the time average of the square of the optical field. By introducing the dielectric constants $\epsilon = 1 + \chi^{(1)}$ and $\epsilon_2 = \chi^{(2)}$ we can combine Eqs. [2] and [3] to get

$$\vec{P} = \epsilon_o [(\epsilon - 1) + \epsilon_2 \langle \vec{E} \cdot \vec{E} \rangle] \vec{E} \qquad [4]$$

where we can now write a total susceptibility as,

$$\chi_t = (\epsilon - 1) + \epsilon_2 \langle \vec{E} \cdot \vec{E} \rangle \qquad [5]$$

or in terms of a total dielectric constant $\chi_t = \epsilon_t - 1$, so that

$$\epsilon_t = \epsilon + \epsilon_2 \langle \vec{E} \cdot \vec{E} \rangle \qquad [6]$$

The corresponding scalar expression for the refractive index can be written as

$$n_t = n + n_2 |E_o|^2 / 2 \qquad [7]$$

for $n_2 \ll n$ where n is determined by $\chi^{(1)}$ and n_2 by $\chi^{(2)}$ in Eq. [1]. This expression is one that explicitly shows the nonlinear dependence of the refractive index, where E_o is the amplitude of the electric field.

The nonlinear dependence of the refractive index may be caused by a number of mechanisms. We are interested in understanding these mechanisms since they lead to the important nonlinear self-actions of focusing and defocusing which may play an important role, and quite possibly the dominant role, in our nonlinear transmission experiments, as well as in a number of other nonlinear experiments.

In the past many of the optical nonlinearities observed in materials, such as two photon absorption, stimulated scattering, and self-focusing imposed an upper limit on the ability of a material to handle increased laser powers. For example, in self-focusing liquids the threshold for stimulated Raman emission depends not so much on the strength of the Raman scattering cross section as on the self-focusing properties of the liquid (7). Self-focusing can also play the dominant role in bistable transmission (8), can give rise to two photon absorption, and dielectric breakdown followed by nonlinear absorption in the laser induced plasma (9).

A number of the mechanisms which give rise to nonlinear refraction have already been identified in solids, liquids, and gases, many of which are organic compounds. A detailed discussion of these mechanisms can be found in a number of good review articles on the subject (10), (11), (12) so they will only be briefly mentioned here. The mechanisms observed to cause nonlinear refraction in organic compounds include: electrostriction, molecular reorientation, molecular redistribution or microscopic clustering, molecular rocking or molecular liberation, electronic polarizability, thermal effects, and resonance effects such as index saturation, Raman scattering, two photon absorption, induced phase transitions, and plasma absorption.

Of the mechanisms mentioned above, molecular reorientation, molecular redistribution, molecular rocking, and electronic polarizability depend only on the local mean-squared field associated with the optical pulse and are rapid compared to nanosecond laser pulse rise times. Each of the above mechanisms leads to a refractive index with a nonlinear dependence on the field associated with the electromagnetic wave and is given by Eq. [7].

Note that Eq. [7] implicitly states that the nonlinear dielectric response time is much slower than the period of the optical oscillations and if we assume that the intensity dependent term in Eq. [7], $n_2|E_0|^2/2$ is always very small compared to the unperturbed refractive index n we can obtain a reasonably simple understanding of self-focusing. Suppose we have a beam of circular cross section with initial beam diameter D which is propagating in a medium of index n. The index in the beam is given by Eq. [7]. A propagating ray of this beam is shown in Fig. [1].

By the use of Snell's law and the condition for total internal reflection $\phi = \pi/2$ we can see that if the angle θ that a propagating ray makes with the propagation direction is less than the critical angle $\theta_c = \pi/2 - \phi'$ then we have met the self-trapping condition and $\cos\theta_c = n/n_t$. From the two approximations $\cos\theta_c \cong 1 - \theta_c^2/2$ and $n_2|E_0|^2/2 << n$ we can write the critical angle for self-trapping as $\theta_c \cong (n_2|E_0|^2/n)^{1/2}$. Now a beam that is limited by diffraction will spread with a half angle given by $\theta \cong 1.22\lambda/2nD$.
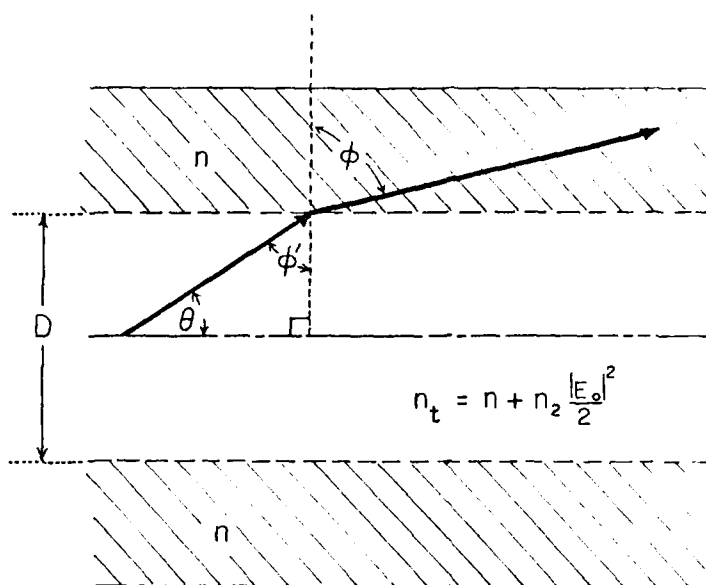
FIGURE I

When $\theta > \theta_c$ the beam spreads by diffraction and when $\theta < \theta_c$ the diffracted rays are totally internally reflected at the boundary of n and n_t and return to the beam. In other words, for propagating Gaussian beams with a cross sectional intensity distribution given by

$$|\vec{E}|^2 = |E_o|^2 \exp [-8r^2/D^2] \qquad [8]$$

the increase in the index in regions of high intensity, i.e., along the axis of propagation gives rise to a lens effect. This comes about since the central portion of the beam experiences a greater phase lag than the off axis portions of the beam causing the intensity along the axis of the beam to increase or in other words to self-focus (13). This increased intensity is accompanied by a reduction in the effective beam diameter which continues to some arbitrary diameter until it is limited by other factors such as saturation effects (14), thus the beam is self-trapped (15).

The critical field for the self-trapping case can be found when $\theta = \theta_c$,

$$|E_o|^2 = \frac{n}{n_2} \left(\frac{1.22\lambda}{2 \, n \, D}\right)^2 \qquad [9]$$

Now the total power in the beam is equal to the integrated intensity.

$$P = \int_o^\infty (n^2 \epsilon_o |\vec{E}|^2/2)(c2\pi r/n)dr \qquad [10]$$

where $|\vec{E}|^2$ is given by Eq. [8]. We can now arrive at a critical power by using Eq. [8], [9], and [10].

$$P_{cr} = \epsilon_o \pi c(1.22\lambda)^2/32n_2 \qquad [11]$$

A more accurate expression for the critical power has been obtained by solving the nonlinear wave equation for a focused beam with a Gaussian intensity profile (16).

$$P_c = 1.86c\lambda^2/16\pi^2 n_2 \qquad [12]$$

where n_2 is the nonlinear index of refraction in ESU units and the conversion from ESU to MKS is given by 9×10^8 (n_2) MKS = (n_2) ESU.

Thus we see that a beam may be trapped at any arbitrary diameter and further that the self-trapping occurs at a critical power level P_c which is independent of the beam diameter. In other words, when the total power carried by a diffraction limited beam exceeds a critical amount, $P > P_c$, then the beam self-focuses and is trapped regardless of the beam size or intensity. When $P < P_c$ the beam diverges due to diffraction, and for the case of $P = P_c$ the beam propagates without convergence or divergence.

These conditions provide a straightforward way to identify self-focusing as a dominant mechanism in nonlinear transmission experiments. The onset of self-focusing is dependent upon power rather than the input intensity and as such the onset of nonlinear transmission would be independent of the focal length of the focusing optics (9). This provides a very reliable way to test if the nonlinear transmission is due to self-focusing rather than two photon absorption or some other mechanism.

EXPERIMENT

The laser system used in these experiments is shown in Fig. [2] and the details of the sampling portion of the setup (enclosed in dashed lines) is shown in Fig. [3]. This system is a passively mode-locked Nd:YAG laser operated at 1.06μm and also at the doubled frequency 0.53μm. It has been described in detail by Van Stryland et al (17); here we only mention its basic features. Single pulses possessing a well defined Gaussian intensity distribution were switched from the pulse train. The temporal pulse width was variable from 30 to 200 picoseconds (FWHM) with the energy in the pulse variable up to 10 mj. Second harmonic autocorrelation scans were used to measure the pulse widths. The pulse widths were monitored after a method described by Glenn and Brienza (18). The energy in the pulse was varied with a polarizing attenuator arrangement and the power in each pulse monitored at the input with an absolutely calibrated detector D_I. A quarter wave plate was used to provide circularly polarized light for some experiments.

The nonlinear transmission experiment is set up after the
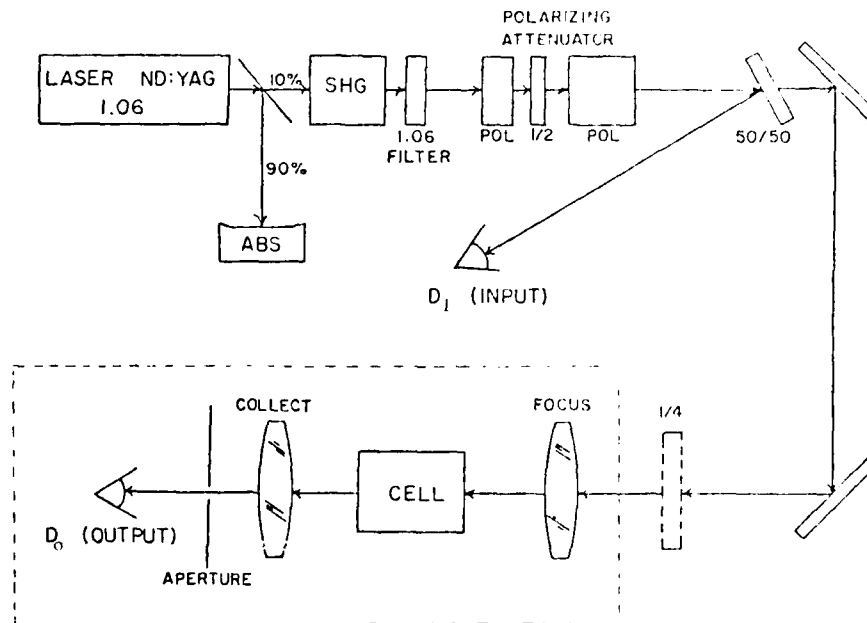
FIGURE 2

arrangement shown in Fig. [3]. A beam having a total power
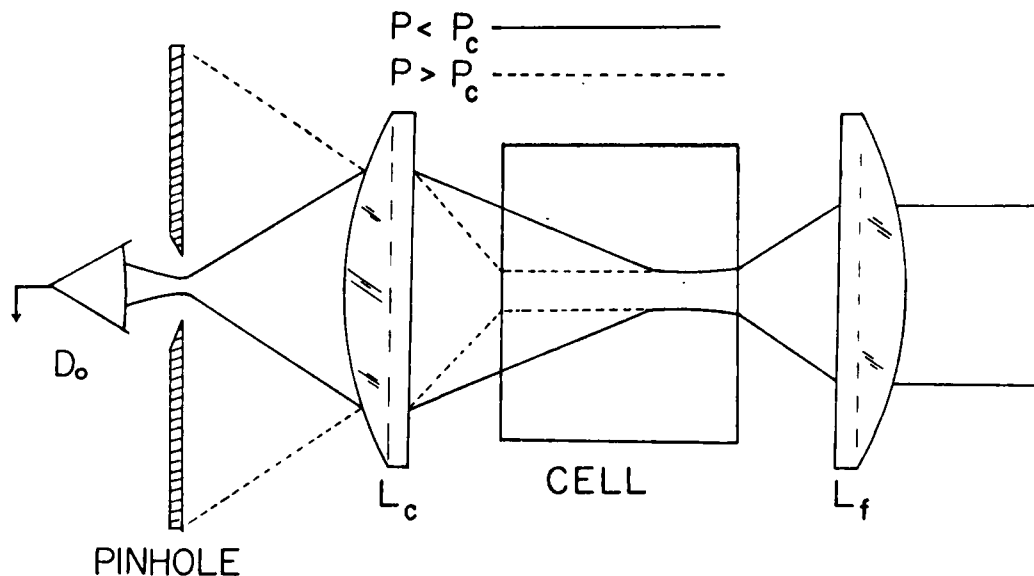
FIGURE 3

less than the critical power $P < P_c$ can be schematically traced as the solid lines in Fig. [3]. For a beam with total power greater than the critical power $P > P_c$ a situation represented by the dashed lines exists for self-focusing. That is, when the input power reaches P_c, the threshold power for self-focusing, the beam suffers severe phase abberations and the apparent focal waist from the focusing lens L_f is moved inside the focal point of the collecting lens L_c and can no longer be imaged through the aperture onto D_o. The threshold power for this is given by Eq. [12] from which we can determine the nonlinear index n_2. It should be noted that the threshold power for nonlinear transmission can be represented by Eq. [12] only when the mechanism responsible for the nonlinear transmission is self-focusing.

In general, the threshold for nonlinear transmission could be due to a number of nonlinear phenomena besides self-focusing, including multiphoton absorption, intensity dependent absorption associated with laser induced breakdown, and stimulated scattering. Thermal effects are ruled out for our experiments due to the low energies and short pulses used. For highly transparent materials such as we studied (see Fig [4], for example, between 0.45μm and 1.10μm).

all of the incident power will be transmitted for input powers be-
low the threshold level, except for the Fresnel reflection losses
at the cell windows. When the threshold level is exceeded and the
transmission of the cell becomes nonlinear we are still able to
distinguish the mechanism responsible by slightly altering our experi-
mental arrangement in Fig. [3]. By removing the aperture in front
of D_o and the collecting optics we can position D_o in close proximity
to the output of the cell such that it intercepts all of the radiation
transmitted by the cell. In this geometry only loses due to absorptive
or scattering processes will be able to change the response of D_o. By
monitoring the intensity dependence of these processes we are able to
identify them and to distinguish them from one another.
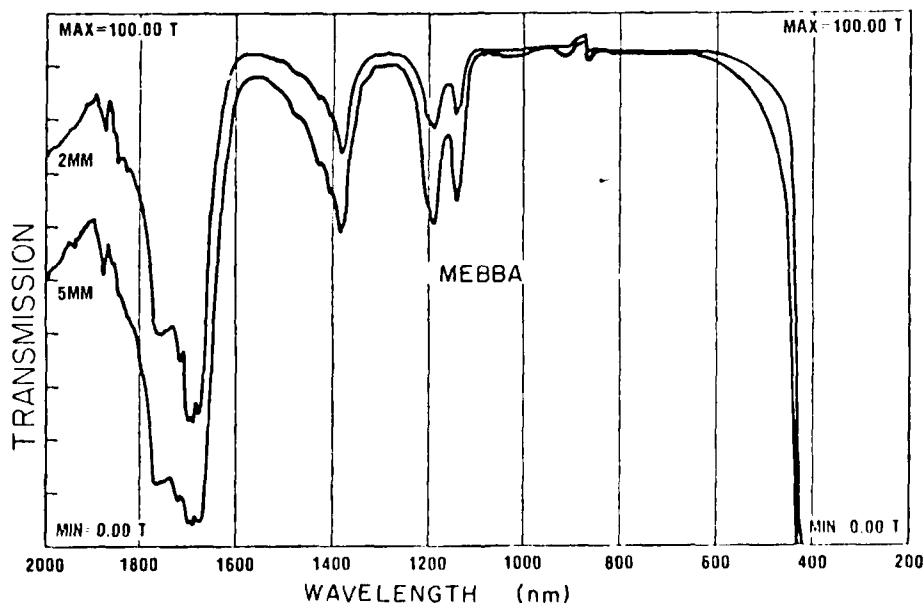
FIGURE 4

MATERIALS

The materials selected for nonlinear transmission measurements
included two chemically similar nematic liquid crystals in the
isotropic state. We observed optical self-action (self-focusing) in

4-methylbenzylidene 4'-n-butylaniline (MEBBA) and in the methoxy
derivative of this compound, 4-methoxybenzylidene 4'-n-butylaniline
(MBBA). Both materials were examined under a variety of conditions
and the threshold powers for self-focusing were determined as a
function of pulsewidth, wavelength, polarization state, and initial
focusing conditions. Figure [5] dipicts the similarity in the

$$CH_3—O—\langle\bigcirc\rangle—CH = N—\langle\bigcirc\rangle—CH_2—CH_2—CH_2—CH_3$$

MBBA

$$CH_3—\langle\bigcirc\rangle—CH = N—\langle\bigcirc\rangle—CH_2—CH_2—CH_2—CH_3$$
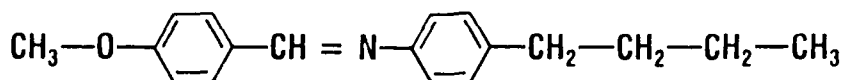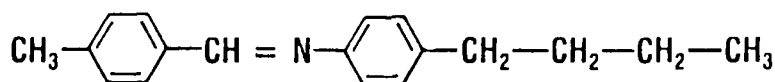
MEBBA
FIGURE 5

chemical structure of these two compounds. Even though these
two materials differ chemically by only a single oxygen atom the
nonlinear optical properties are significantly different. The
measured threshold powers for self-focusing in MBBA and MEBBA
differ by an order of magnitude. Finally we compare the results
obtained on these two materials with the well known self-focusing
liquid carbon disulfide CS_2.

RESULTS

Using the experimental arrangement shown in FIG. [3] with a
75mm focal length lens to provide the initial focusing, we recorded
the detector response D_o as a function of the input power for MEBBA
and MBBA. Figure [6] shows the results of these measurements for
22 ± 5 ps pulses at 0.53µm for the linear polarization state. Each
data point on these curves represent the average value of D_o in
arbitrary units for 5 individual pulses which were selected from
the microprocessor controlled ND:YAG laser system. The 5 pulses
selected for averaging were picked for their uniformity to pre-
determined pulse width and energy content. Any pulses which fell
outside the predetermined windows for the pulse width and energy
were rejected and the D_o value recorded for these were not used
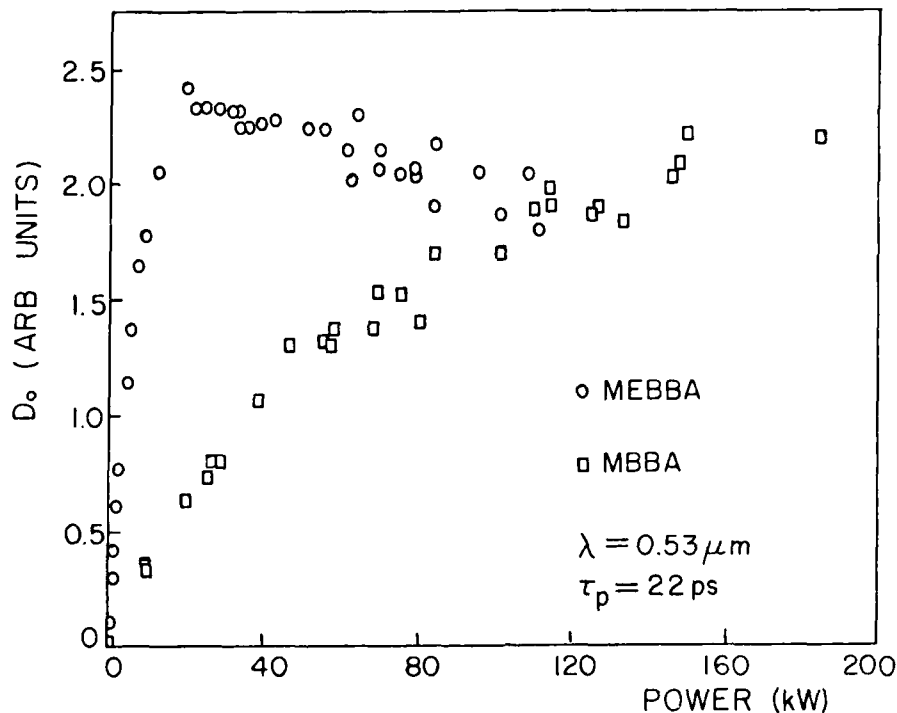
FIGURE   6

in the averaging process. This procedure minimized the random errors
which could be injected by variations in pulse width and energy content
from pulse to pulse. The behavior of these two materials in the region
of the threshold are clearly not the same since MBBA reaches a saturation
value more gradually than MEBBA. The lack of a defined turnover
point in the curve indicates that mechanisms other than self-focusing
may add to the nonlinear transmission such as multiphoton absorp-
tion for example. The exact value of the threshold power is therefore
not as clearly determined for MBBA as it is for MEBBA but the order
of magnitude difference in the threshold power is easily discernable
for these conditions. These measurements were repeated for circularly
polarized light and are the same within experimental error indicating
that there is no polarization dependence for these two materials.

The mechanism reponsible for the onset of nonlinear transmission
in MEBBA is self-focusing since we observed the same threshold powers
for different initial focusing conditions. Figure [7] shows the non-
linear transmission for two different focal length lenses which were
used to focus the light into the sample cell. The behavior near the
onset of nonlinear transmission is different for these two cases but
the threshold powers are the same within experimental error.
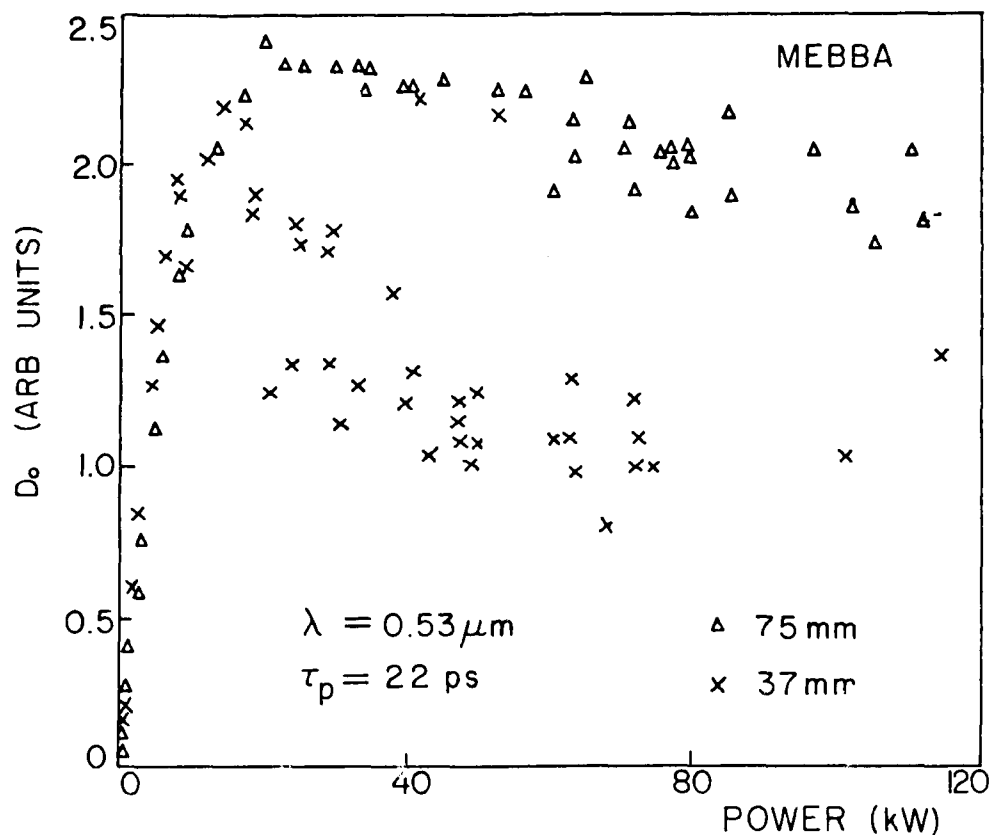
FIGURE 7

This indicates that the mechanism responsible for the onset of the nonlinear transmission depends only on the total beam power. This is a strong indication that self-focusing is the mechanism dominating nonlinear transmission and therefore the threshold power can be described by Eq. [12].

The relatively large variations in the data after the onset of nonlinear transmission for the 37mm lens case is probably due to absorption by laser induced plasmas which were observed in the liquid during the measurements. They were not observed for the 75mm lens, so for a given power they were probably formed by the higher irradiance levels produced by the shorter focal length lens.

In Fig. [8] we show the nonlinear transmission data for MEBBA taken with 0.53μm linearly polarized pulses at 22 $\pm$ 5ps and 112 $\pm$ 22ps. We compare it to CS_2 for the same wavelength and polarization at 22 $\pm$ 5ps. Soileau et al (9) have shown that the threshold power for CS_2 is independent of pulse width from approximately

62
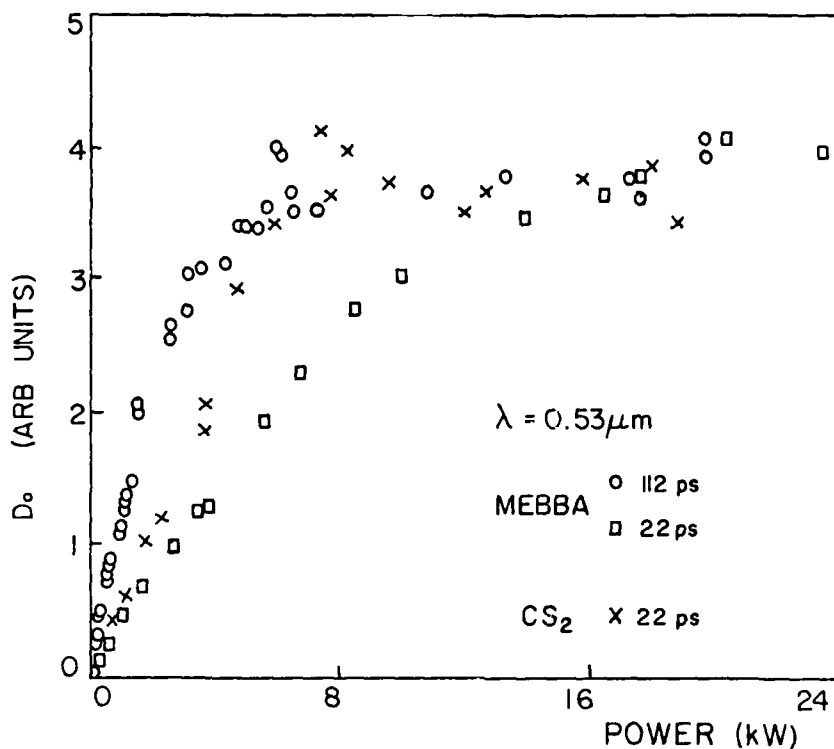
FIGURE 8

40ps to 9ns and since the relaxation time of CS_2 is approximately 2ps (19), (20) we expect the threshold power for CS_2 to be nearly independent of pulse width in the 22 to 112ps range. Recent measurements (21) of the nonlinear index of CS_2 using a new direct technique produced a value of $n_2 = 1.23 \times 10^{-11}$ ESU for approximately 30ps pulses. Using Eq [12] and this value of n_2 at 0.53μm we calculate a threshold power of 7.8kW which agrees very well with the data of Fig. [8].

There is however a considerable pulse width dependence observed for MEBBA with the value of the threshold power changing from $\sim$ 12kW at 22ps to about 2.5kW at 112ps and if extrapolated to the nanosecond regime we might obtain substantially lower values of P_c than those observed here for 0.53μm pulses. Such an extrapolation is not un-realistic if we examine available data on MBBA. Rao, et al (22) observed self-focusing in MBBA for 20ns ruby pulses at 0.69μm. They report a value of $n_2 = 2.3 \times 10^{-10}$ ESU for these conditions and we

63

get a value of $n_2 = 1.26 \times 10^{-12}$ ESU for 22ps pulses at 0.53μm or nearly two orders of magnitude difference.

We made similar measurements on MEBBA at 1.06μm and these showed that the onset of nonlinear transmission occurred at powers approximately 70 times higher at 1.06μm compared to 0.53μm, which implies that there is considerable dispersion in n_2 in the pulse length range of these experiments.

Finally for comparison we compute the nonlinear index for MEBBA showing threshold powers of 11kW associated with 22ps linearily polarized pulses at 0.53μm to be $n_2 = 8.98 \times 10^{-12}$ ESU and for 2.5kW associated with 112 ps, linearly polarized pulses at 0.53μm to be $n_2 = 3.95 \times 10^{-11}$ ESU.

This strong pulse width dependence observed in MEBBA for ps pulses rules out bound electronic effects (fast compared to picoseconds) as a mechanism for the observed self-focusing. The fact that there is a large dispersion in n_2 for ps pulses implies that the mechanism is not simply molecular reorientation as it is in CS_2. We suggest that two possible mechanisms for the observed nonlinearity are bond switching and saturation.

The relatively large and relatively fast nonlinearity exhibited by MEBBA could have application in a wide range of new optical devices. The nonlinear properties of organic materials have already been used in a number of ways to condition the intensity and spectral contant of optical radiation. Some of these uses include such phenomena as: fast optical gating (20), (23), ultrafast two dimensional photography (24), second harmonic generation (25), stimulated scattering (7), (14), degenerate 4-wave mixing (26), wavefront reversal (27), optical bi-stability (8), (28) and finally self-focusing (8), (22), (29) which is intimately connected to many of the above nonlinear effects. Nearly all of these phenomena have been studied in liquid crystals and often the nonlinearity is considerable (4). The results presented here on MEBBA serve to strengthen the contention that organic materials will play an ever increasing role in the rapidly growing area of nonlinear optics and in applications stemming from their use.

ACKNOWLEDGEMENTS

REFERENCES

1. Hochstrasser, R. M., and Trommsdorff, H. P., Acc. Chem. Res. 16:376 (1983).

2. Garito, A. F., and Singer, K. D., Laser Focus pp. 59-64, Feb. (1982).
3. Lalama, S. J., Singer, K. D., Garito, A. F., and Desai, K. N., Appl. Phys. Lett. 39:940 (1981).
4. Ya Zel'dovich, B., Pilipetskii, N. F., Sukhov, A. V., and Tabiryan, N. V., JETP Lett. 31:263 (1980).
5. Khoo, I. C., and Zhuang, S. L., Appl. Phys. Lett., 37:3 (1980).
6. Sharp, E. J., "Electro-Optics and Nonlinear Optics" in Electronic Engineers Handbook, (Eds. D. Fink and D. Christiansen) McGraw Hill; New York, (1982).
7. Wang, C. C., Phys. Rev. Lett. 16:344 (1966).
8. Khoo, I. C., Zhuang S. L., and Shepard, S., Appl. Phys. Lett. 39:937 (1981).
9. Soileau, M. J., Williams, W. E., and Van Stryland, E. W., IEEE J. Quantum Electron., QE-19:731 (1983).
10. Shen, Y. R., Progress of Quantum Electronics, (Eds. J. H. Sanders and S. Stenholm) Pergamon; New York, (1977).
11. Kielich, S., IEEE J. Quantum Electron. QE-4:744 (1968).
12. Kerr, L. E., Phys. Rev. A, 4:1195 (1971).
13. Kelley, P. L., Phys. Rev. Lett. 15:1005 (1965).
14. Rahn, O. and Maier, M., Phys. Rev. Lett. 29:558 (1972).
15. Chiao, R. Y., Garmire, E., and Townes, C. H., Phys. Rev. Lett. 13:479 (1964) and Erratum, ibid. 14:1056 (1965).
16. Marburger, J. H., Progress of Quantum Electronics, (Eds. J. H. Sanders and J. Stenholm) Pergamon: New York, (1977).
17. Van Stryland, E. W., Soileau, M. J., Smirl, A. L., and Williams, W. E., Phys. Rev. B. 23:2144 (1981).
18. Glenn, W. H. and Brienza, M. J., Appl. Phys. Lett. 10:221 (1967).
19. Reintyis, J., Carman, R. L., and Shinizu, F., Phys. Rev. 48:1486 (1973).
20. Ippen, E. P. and Shank, C. V., Appl. Phys. Lett. 26:92 (1975).
21. Williams, W. E., Soileau, M. J., and Van Stryland, E. W., "Fifteen Annual Symposium on Optical Materials for High Power Lasers," Boulder, CO, Nov 1983 (to be published).
22. Rao, D. V. G. L., Narasimha, and Jayaraman, S., Appl. Phys. Lett. 23:539 (1973).
23. Etchepare, J., Grillan, G., Migus, A., Martin, J.L., and Hamoniaux, G., Appl. Phys. Lett. 43:406 (1983).
24. Sala, K. and Richardson, M. C., Phys. Rev. A. 12:1036 (1975).
25. McFee, J. H., Bergman, J. G. Jr., and Crane, G. R., Ferroelectric, 3:305 (1972).
26. Khoo, I. C., Phys. Rev. A, 25·1040 (1982).
27. Khoo, I. C. and Zhuang, S. L., IEEE J. Quantum Electron. QE-18:264 (1982).
28. Liao, C., and Stegeman, G. I., Appl. Phys. Lett. 44:164 (1984).
29. Wong, G. K. L. and Shen, Y. R., Phys. Rev. Lett., 32:527 (1974).

A QUANTITATIVE METHOD FOR DETERMINING
THE EXTENT OF SKIN IRRITATION

MICHAEL C. POWANDA, Ph.D., LTC, MSC, PETER SCHMID, Ph.D.
JUNE JAEGER, M.S., GEORGE J. KLAIN, Ph.D. & JOHN P. HANNON, Ph.D.
DIVISIONS OF CUTANEOUS HAZARDS & COMBAT CASUALTY CARE
LETTERMAN ARMY INSTITUTE OF RESEARCH
PRESIDIO OF SAN FRANCISCO, CA 94129

Vesicants, in particular sulfur mustard, have long been considered to be an exceedingly effective means of denying territory and materials to combatants and support personnel who are not prepared to protect themselves against, or to decontaminate themselves after being exposed to, this class of chemical warfare agents. Sulfur mustards are especially insidious in that decontamination is unlikely to be effective unless applied within 10-15 minutes of exposure to the vesicant (1), though recent evidence indicates that changes in cutaneous cellular metabolism occur within 1 minute of exposure to sulfur mustard (2). Thus, one essential aspect of the chemical defense program requires that we are able to rank order the effectiveness of formulations thought to protect against and, if possible, decontaminate the skin of personnel exposed to such agents. In addition, we need to insure that the formulations designed to protect and/or decontaminate the skin against vesicants and other chemical warfare agents, such as organophosphates, are not of themselves irritants. This is to preclude fielding systems, such as the various forms of the M258 kit, which can themselves cause pronounced skin irritation (3). Hence, these systems cannot be used readily in training and, moreover, if improperly used under combat conditions, they could enhance rather than reduce the problem of skin decontamination.

The objectives of these studies are twofold, (a) to determine the feasibility of developing a system of quantitative measures of edema and erythema to provide an assay system for assessing the efficacy of formulations that protect against vesicants, particularly sulfur mustards; and (b) to determine whether such a system could have a general application to the problem of testing topical protection/ decontamination formulations for use against organophosphates, as well as vesicants, for intrinsic dermal irritancy.

The various procedures presently used for determining skin

irritation employ the Draize scoring system developed in 1944 (4).
This system requires visual inspection of the shaved skin of an animal,
usually a rabbit, to determine the extent of erythema (abnormal redness
of skin) and edema (leakage of serum protein containing fluid). This
dependence upon subjective visual inspection to determine the extent
erythema and edema is the major reason for considerable inter-
laboratory differences in classifying and rank ordering irritants
(4,5). We, therefore, have undertaken to develop a system which will
allow contractors and in-house investigators to quantitate objectively,
in a user-independent fashion, the physiologic responses to skin
irritation.

We have chosen the rabbit as the animal component of the test
system because it often shows a stronger response to mild and moderate
irritants than humans, thereby providing an additional margin of safety
(6). Since edema is the result of protein-containing fluid
infiltration into the site of irritation, we measured both the water
content of tissue and the accumulation of albumin labelled with
radioactive iodine, atomic number 125 (^{125}I-albumin), at the site of
irritation (7). Albumin is the major protein component of such
effusions. As an index of erythema, we first measured the accumulation
of red blood cells tagged with radioactive chromium atomic number 51
(^{51}Cr-erythocytes) and then switched to laser Doppler velocimetry when
the equipment arrived. The irritant chosen was beta chlorethyl n-butyl
sulfide, a potent vesicant and analog of the sulfur mustard agent, HD.
Glass cups containing filter paper impregnated with vesicant (Figure
1a) were attached by means of double-stick discs to the shaved dorsum
of an anesthetized rabbit for 30 minutes (Figure 1b). At three hours
^{125}I-albumin and ^{51}Cr-erythrocytes were injected intraveneously and
allowed to circulate and equilibrate for one hour. At four hours, a
sample of blood was taken, the rabbit killed with an overdose of T-61
euthanasia solution (American Hoechst Corp) and a large portion of the
dorsal skin removed and mounted on cardboard to reduce shrinkage
(Figure 1c). The sites of application of vesicant and the control
site(s) were excised with a rotating cork borer (Figure 1d). These
tissue samples were vacuum-dried to determine the water content and the
dried samples placed in glass vials, dissolved in 30% aequeous
potassium hydroxide at 90°C and counted in a dual-Channel gamma counter
to assess accumulation of ^{125}I and ^{51}Cr.

Table 1 presents the data and coefficients of variation of the
tissue water response from control sites or sites exposed to 10 ul of
vesicant for 30 minutes. The data are presented so one can evaluate
animal and/or experiment, as well as, site variation, and perceive the
effect of vesicants on the variables measured. In order to allow
estimation of site variability, the dorsal area of the rabbit was
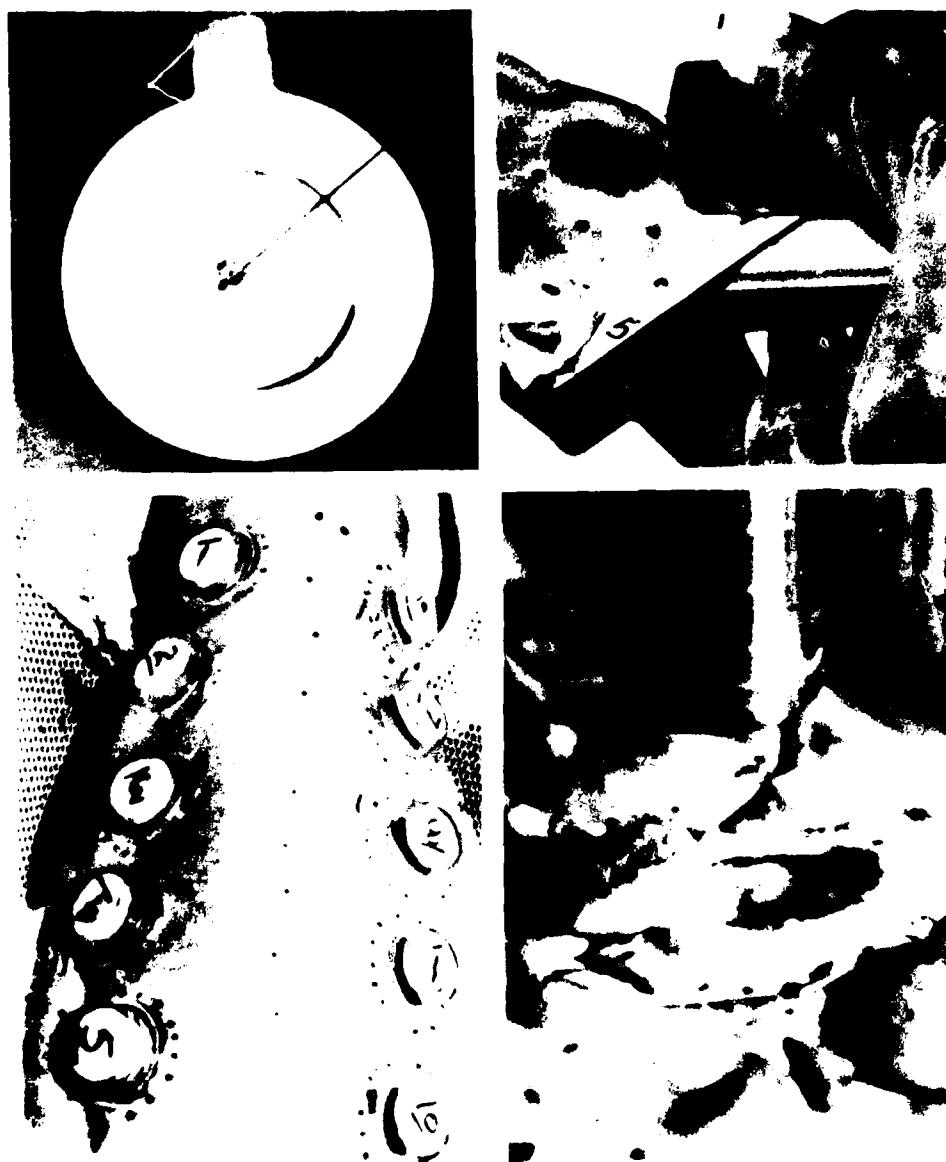
Figure 1. (a) top left, glass cap containing vesicant; (b) bottom left, caps attached to dorsal aspect of an anesthetized rabbit; (c) top right, excised skin attached to cardboard mount; (d) bottom right, skin sample removed with cork borer.

divided into 10 sites, 5 on each side of the spine. Vesicant was applied to one side, the other side provided control samples, thereby generating contralateral, paired sites to minimize variation and facilitate data analysis.

TABLE 1: Percent Water

		Experiment Number (Animal/Experiment Variation)					
		10	14	15	16	17	18
Vesicant	Mean	78.0	77.9	81.6	81.3	80.4	84.4
	S.D.	3.1	1.8	1.6	0.3	1.2	1.7
	C.V.	4.0%	2.3%	1.9%	0.4%	1.5%	2.0%
Control	Mean	66.0	70.9	70.2	72.9	72.5	73.1
	S.D.	1.6	1.4	3.1	0.9	1.1	0.6
	C.V.	2.4%	2.0%	4.4%	1.2%	1.5%	0.8%

		Site Number (Site variation)				
		1	2	3	4	5
Vesicant	Mean	80.9	81.0	81.4	79.5	80.6
	S.D.	0.9	2.1	3.0	4.0	3.5
	C.V.	1.2%	2.7%	3.7%	5.1%	4.4%
		6	7	8	9	10
Control	Mean	72.2	71.6	70.9	70.5	69.5
	S.D.	2.6	2.6	2.1	2.6	4.2
	C.V.	3.6%	3.7%	2.9%	3.7%	6.1%

The top half of Table 1 gives a sense of the variation between animals; the mean, S.D. and C.V. for 5 sites are displayed. The mean % water at the control site varied from 66 to 73% while the % water at the sites of vesicant application varies from 77 to 84% indicating interanimal/experiment variation that may be susceptible to reduction. In contrast, when the data for the same site were averaged from animal to animal as shown on the bottom half the table, there was little variation amongst sites.

Table 2 presents the cpm ^{125}I-albumin data. Vesicant induces a 3 to 9-fold increase in accumulation of radiolabelled albumin at the exposure sites. There is evident inter-and intra-animal variation which may indicate that equilibration of the labelled albumin has not yet occurred. Nonetheless, when one assesses site variation, the range of mean values for vesicant treated sites is narrow (620-661 cpm) as is that for control sites (95-109cpm).

TABLE 2:   CPM ^{125}I-Albumin/Sample

		Experiment Number (Animal/Experiment variation)				
		10	14	16	17	18
Vesicant	Mean	731	428	444	593	990
	S.D.	127	99	53	133	143
	C.V.	17%	23%	12%	22.5%	14.4%
Control	Mean	96	131	72	107	106
	S.D.	15	11	8	20	16
	C.V.	16%	8.1%	12%	19.2%	15.4%

Negligible site variation indicates that on any given rabbit, there is no significant difference in response as long as the chemicals are applied on the shaved dorsal surface. The animal/ experiment variation may be due to differences in stages of hair growth cycle, which can affect capillary blood flow, or differences in skin and ambient temperature which can also affect capillary blood flow, as well as individual animal differences. Some of these variations can be minimized or at least controlled.

There were no alterations in ^{51}Cr deposition in vesicant treated sites despite the presence of erythema. A recent review of the literature notes that other investigators also have found that the labelling of erythrocytes with ^{51}Cr does not provide a sensitive measure of erythema (8). Thus, an alternative approach was taken to quantitate erythema, but that depended on the acquisition of equipment which has only recently arrived. These data will be displayed later.

In the absence of skin protection formulations to test, we evaluated the relative protection afforded by a single thickness of battle dress uniform (BDU) fabric versus Gortex$^{(R)}$ against half-sulfur mustard vapor. Three swatches of each cloth were placed between the cups containing vesicant and the rabbit's skin for the duration of the exposure (Figure 2). As a positive control, the skin was exposed to vesicant with no cloth intervening. The mean and S.D. for two experiments are shown (Table 3). The somewhat larger than anticipated variation is ^{125}I-albumin values may be due to the decay of the radiolabel and the lower counts observed in the 2nd study. Despite limited data, it can be seen that both BDU fabric and Gortex$^{(R)}$ provide some protection against mustard vapor. Conversely, these data also demonstrate that this system can be used to evaluate not only formulations, but also clothing, for their ability to protect against vesicant. These data also emphasize the value of measuring both tissue water content and ^{125}I-albumin to obtain the most conservative

71

Figure 2. Left, vesicant cups with BDU (top) and Gortex$^{(R)}$ swatches attached; right, vesicant cups attached to dorsal aspect of an anesthetized rabbit.

evaluation. Based on ^{125}I-albumin accumulation, it appeared that both materials afforded nearly complete protection while the tissue water content suggested the protection was not complete.

TABLE 3. Protection by Fabric Against Half-sulfur Mustard Vapor.

Group	% Tissue Water[a]	CPM ^{125}I-Albumin/ Sample[a]
Control	72.0	153
	1.4	98
Vesicant Alone	79.9	600
	3.2	205
Vesicant + BDU	77.4	249
.55mm thick	2.8	36
Vesicant + Gortex[(R)]	75.8	169
.40mm thick	2.4	54

30 minute vesicant exposure, 10 ul vesicant. The tissue samples were taken at 4 hours, as in preceding studies.

[a] mean and S.D., n=4 for control, 6 for experimental sites.

If Gortex$^{(R)}$ also markedly reduces the penetration of nerve agents, as measured in the in vitro test system described by Hawkins and Reifenrath (9), or can be impregnated with an enzyme capable of degrading organophosphates akin to the DFPase reported by White et al., (10), then it may be possible to field a light weight, water repellent, protective uniform which does not cause appreciable heat-induced fatigue in troops.

The Army is presently asking contractors to develop a long-lasting ᵃ single application effective up to 12 hours), user-acceptable insect repellent to protect soldiers from insect vectors of disease and reduce personnel losses. It is essential, considering the likelihood of chemical warfare, to ascertain whether such a product will enhance the possibility of chemical casualties by facilitating the transport of such chemicals through skin. This test system also will be able to answer that question with respect to vesicants.

As indicated earlier, the use of ^{51}Cr-labelled erythrocytes was

not effective in assessing the erythema which developed in response to vesicant exposure. We thus began studies to determine if laser Doppler flowmetry, which has been used to provide a non-invasive method of measuring blood flow through skin (11), would also provide a measure of erythema. The basic concept is that the frequency of laser light (632.8nm) back scattered from moving red blood cells is shifted in frequency in proportion to the velocity of the red blood cells. The actual value that one obtains from the use of commercially available laser Doppler flowmeters also reflects the number of red cells present, thus, it seemed reasonable to apply this technique to the problem at hand.

A laser Doppler flowmeter was obtained from Med Pacific Corp, Seattle, Washington. The instrument came with an optical fiber attachment which allows one to direct the light to the tissue site and return the reflected signal to the instrument for processing (Figure 3a). So as not to contaminate the probe, we did not begin taking measurements until 30 minutes after the half exposure to vesicant was completed. The data from two rabbits is shown as a composite in Figure 3b. The output of the instruments is given in millivolts, the larger the value, the greater the blood flow and/or concentration. The values shown are the sum of 5 control and 5 experimental sites on each animal. Within 30 minutes after exposure to vesicant was completed, the experimental sites display an almost two-fold increase in signal; this increased signal persists for at least 5 hours and is still somewhat detectable at 24 hours after vesicant exposure (data not shown).

In order to obtain reproducible values, the animals must be restrained or anesthetized to minimize movement; in this study, we used ketamine combined with acepromazine to sedate the animals. When our controlled temperature chamber is built, we will test the effect of alterations in ambient temperature on blood flow, as well as on the tissue response to vesicant. Based on systemic responses to inflammation, it may be necessary to maintain the rabbits at their thermal neutral zone throughout the period of study.

SUMMARY

1. We have applied basic physiologic knowledge and 20th century technology to the problem of quantitating edema and erythema and as a result, an in vivo test system has been developed which will allow the quantitative, user-independent assessment of topical formulations designed to protect against vesicant-induced skin injury/irritation.

2. Some components of the test system need to be optimized to reduce further the likelihood of interlaboratory variation in results. The

LASER LIGHT PATH SCHEMATIC
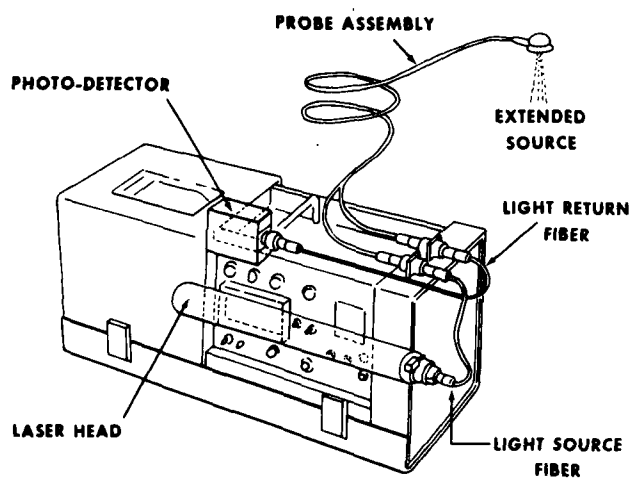
LASER DOPPLER FLOWMETRY
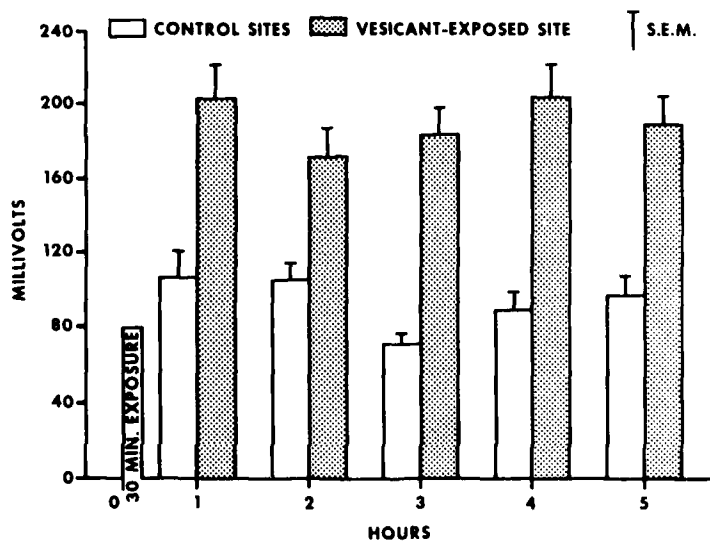
Figure 3.   (a) top, diagram of laser light path;
(b) bottom, data from 2 rabbits, 5 control and 5
vesicant exposed sites each.

75

system should be validated against neat agent, as soon as optimization is accomplished. While validation is ongoing, this methodology can be made available to investigators and contractors so they can evaluate the effectiveness of anti-vesicant formulations.

3.   The system can be readily modified to allow testing of materials, uniforms, and protective gear for their effectiveness in protecting against or enhancing the action of vesicants.

4.   The next phase will be to evaluate the test systems ability to assess the intrinsic dermal irritancy of topical formulations. This should be readily accomplished since we are merely providing quantitative measures of edema and erythema, the two components of inflammation that have been used for over 40 years to estimate the skin irritancy of materials.

5.   Application of this test system to the general problem of dermal irritancy will not eliminate the use of animals to assess the safety of materials likely to ome in contact with skin, but should reduce the number of animals needed to determine the degree of irritancy and diminish interlaboratory variation.

ACKNOWLEDGEMENTS

1. Smith HW, Clowes GHA, and Marshall EK. On Dichloroethylsulfide (mustard gas). IV. The Mechanism of Absorption by the Skin. J. Pharmacol. 1919;13:1-30.

2. Klain GJ, Powanda MC, Schmid P and Black KE. Metabolic Markers of Chemically-Induced Cutaneous Irritation/Injury. Army Science Conference, 1984.

3. Jederberg WW, Fruin JT, Hanes MA. Primary Dermal Irritation Potential of Components of the M-258A-1 Decontamination Kit. Letterman Army Institute of Research, Report No. 120.

4. Draize JH, Woodward G, Calvery HO. Methods for the Study of Irritation and Toxicity of Substances Applied Topically to the Skin and Mucous Membranes. J. Pharm. Exp. Therap. 1944;82:377-419.

5. Weil CS, Scala RA. Study of Intra and Interlaboratory Variability in Results of Rabbit Eye and Skin Irritation Tests. Toxicol. App. Pharmacol. 1971;19:276-360.

6. Dermatotoxicity: Selected Issues for Dermal Toxicity, Including Irritation, Sensitization, Phototoxicity and Systemic Toxicity. Environmental Protection Agency, Report Number 560/11-82-002, Pg 9.

7. Udaka K, Takeuchi Y, Movat HZ. Simple Method for Quantitation of enhanced vascular permeability. Proc. Soc. Exp. Biol. Med. 1970;133:1384-1387.

8. Issekutz AC. Quantitation of Acute Inflammation in the Skin: Recent Methodological Advances and Their Application to the Study of Inflammatory Reactions. Surv. Synth. Path. Res. 1983;1:89-110.

9. Hawkins GS, Reifenrath WG. An In Vitro Method for Determining the Fate of Chemicals Applied to the Skin. Army Science Conference, 1984.

10. White CT, Martin EJ, Unruh KA. A Stable, High Activity Preparation of DFPase for Use in Detection of, and Protection Against, Nerve Agents. Army Science Conference, 1984.

11. Nilsson GE, Tenland T, Oberg PA. Evaluation of a Laser Doppler Flowmeter for Measurement of Tissue Blood Flow TEEE Transactions on Biomedical Engineering. BME 1980;27:597-604.
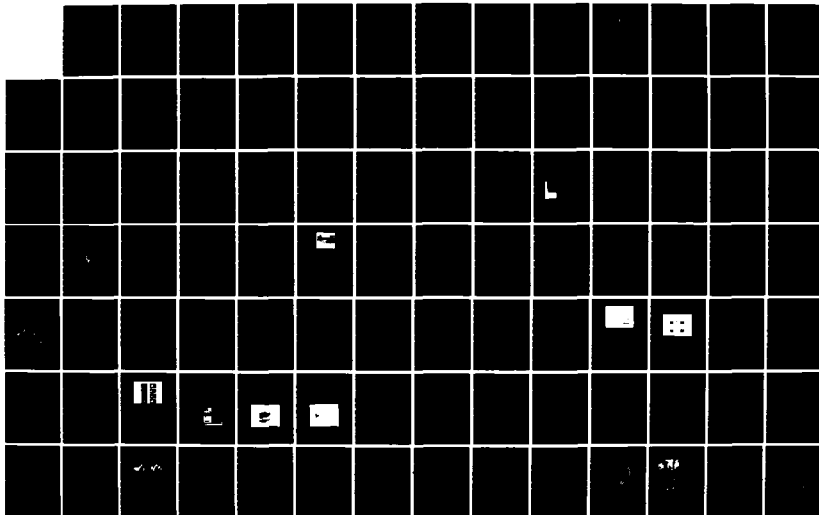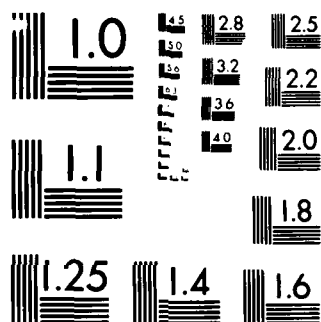
MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

SUB-SURFACE RESIDUAL STRESS MEASUREMENTS IN DEPLETED
URANIUM BY MEANS OF NEUTRON DIFFRACTION (U)

*HENRY J. PRASK, DR.[+]
CHANG S. CHOI, DR.[+]
ENERGETIC MATERIALS DIV., LCWSL, ARDC
U.S. ARMY ARMAMENT, MUNITIONS, AND CHEMICAL COMMAND
DOVER, NEW JERSEY 07801

INTRODUCTION

In recent years depleted uranium (DU) has become increasingly
important in armament-system applications, specifically in armor-piercing
projectiles. It is expected that the next several years will see increased
use of DU in such munitions, with new systems already in the design or R&D
phases. The intrinsic properties of DU are so appropriate for these
applications that increased utilization is expected in spite of the fact
that some very serious gaps have persisted in the ability to characterize
certain critical materials properties. One of the most important of these
is residual stress.

Macro-residual stresses are a distribution of stresses retained within
a body after external stimuli have been removed. They arise, for example,
from nonuniform heat treatment or plastic deformation and vary continuously
in the body over regions which are very large compared to inter-atomic
distances. With regard to performance or durability of military end-items,
the net stress on a body in service is the superposition of external forces
and the residual stress system within the body. This has enormous
importance in fatigue failure, fracture, and in stress-corrosion cracking,
and, therefore, in process design and structural performance. For most
metals, the effect of detrimental residual stresses can be minimized by
stress-relieving heat treatments. In the case of DU, this is not a viable
procedure, as discussed below.

DU-0.75 wt% Ti has become the principal constituent in armor-piercing
projectiles because of its high density, hardness, ductility, and high
ultimate tensile and yield strengths. However, these properties are
achieved by a fabrication process which involves a fast quench to ambient

---

[+]Guest scientists at the Reactor Radiation Division, National Bureau
of Standards, Washington, DC.

temperature from the γ-phase (~ 800 °C), followed by aging at ~ 400 °C. This yields optimal mechanical properties with the Ti "dissolved" in the single-phase α'-U host lattice. Mechanical or heat-produced stress-relief is ineffective until levels are reached where the α'→β phase transition or delitorious plastic deformation occur. Therefore, the (by-default) approach, in principle, has been to attempt to adjust processing procedures to attain optimum residual stress distributions and maximum performance. In practice, enormous difficulties have been encountered with this approach because - as it has now been found - DU presents unique, unsolved problems for residual stress characterization for virtually all conventional techniques.

X-ray diffraction for residual stress analysis (XRDRSA) has been developed so successfully that it is now a routine method for nondestructive determination of surface stresses in metals. In fact, in the last few years portable devices for field use have become commercially available. However when applied to DU, grain size, surface oxidation, and other effects have produced insurmountable problems for XRDRSA, as reported in references (ˆ) and (2). Alternative nondestructive techniques - such as ultrasonics, eddy current, etc. - have seen only limited success on other metals and have not proven useful for DU. Until now, the technique which has been used with the greatest degree of success is the (destructive) Sach's boring-out method (2). Applied to DU, it is slow, expensive, requires special surface coating, special boring equipment, and the use of strain gauges. In addition, rather drastic assumptions are required to extract stress values for the cylindrical specimens. For example, hoop and axial stresses are assumed to vary only with radial distance from the cylinder axis. The fact that the DU bars are bowed when removed from the quench bath shows this to be incorrect.

Neutron diffraction closely parallels x-ray diffraction in methodology and analytical formalism. However, because neutrons interact with nuclei and x-rays with electrons, neutrons are nondestructive and typically about a thousand times more penetrating than x-rays in the wavelength range for diffraction (0.7 Å $\leq \lambda \leq$ 4 Å). In addition, different elements exhibit significantly different relative scattering powers for neutrons and x-rays. A utilization of the unique aspects of neutron diffraction for residual stress analysis was suggested and partially demonstrated previously (3). Recently, neutron diffraction has been used to measure residual stress in certain samples (reviewed in reference 4). However, the energy-dispersive neutron diffraction technique described in this paper and reference 4 minimizes texture and absorption effects which limits conventional measurements. The technique and its successful application to DU are described in the next sections.

EXPERIMENTAL

1. General considerations

In both x-ray and neutron diffraction determination of residual stress, what is measured is strain which is manifested by changes in the distances between atomic planes, or d-spacings, in the sample. A unique advantage of neutron diffraction arises from the different relative scattering cross-sections and penetration relative to x-rays. This is illustrated in Table I in which $t_{1/2}$, the thickness at which half the beam intensity is lost through scattering and absorption processes, is listed for some metals of technological importance. The values are based on cross-sections from standard references and the difference in wavelengths used for neutrons and x-rays is not significant.

Table I. X-ray/Neutron Comparison

Element (Atomic Number)	$t_{1/2}$ (1.54 Å x-rays)	$t_{1/2}$ (1.80 Å Neutrons)
Al (13)	0.0530 mm	71.0 mm
Ti (22)	0.0076	15.9
Fe (26)	0.0027	6.1
W (74)	0.0021	6.5
U^{238} (92)	0.0015	13.6

It should be mentioned that the $t_{1/2}$ values in Table I do not represent the depth of penetration in a residual stress measurement. This is dependent on a number of factors such as source intensity, coherent scattering cross-section, and beam spot size. However, the $t_{1/2}$ values clearly show that neutrons in the normal diffraction wavelength range are several orders-of-magnitude more penetrating than x-rays. Also, the penetration does not decrease monotonically with atomic number as with x-rays, but is, essentially, a random function of Z. It is of particular importance that DU is an especially good material for neutron examination.

The properties of neutrons suggested the possibility of measuring sub-surface residual stresses in metallurgical samples employing tight collimation and the scattering geometry shown schematically in Figure 1. A scattering angle, Ω, of 90° is used to minimize the examined differential volume, ΔV. The perspective view in Figure 1 shows ΔV defined by two rectangular apertures in an absorbing material, e.g., cadmium; however, the apertures could be any shape including circular. The plan view in Figure 1 indicates how the sample can be translated in the beam so that ΔV can be examined as a function of depth.

The strains from which residual stresses are inferred are obtained from measured d-spacings through Bragg's Law: $\lambda = 2d(hkl)\sin \Omega/2$ where λ is wavelength, $d(hkl)$ is the separation of atomic planes with Miller

indices hkl, and Ω is the scattering angle. In a typical x-ray or neutron measurement, λ is fixed and Ω is stepped--usually with the sample orientation stepped by $\Omega/2$--and sharp resonances in scattered intensity are observed at scattering angles where the Bragg condition is fulfilled. Precise determination of $\Omega(hkl)$, the peak position, yields d(hkl) directly. Although some success has been achieved with this mode of measurement, as reviewed in reference 4, serious problems have been encountered with highly attenuating or highly textured samples (5). Here, texture means the existence of preferred crystallographic orientation of grains or crystallites relative to a coordinate system fixed in the sample.

With reference to the plan view in Figure 1, it is clear that as Ω is varied, path length to and from ΔV changes and intensity as a function of Ω is distorted leading to a false shift in d(hkl). Similarly, gradients in preferred grain orientation in the sample over the changing beam-in/beam-out paths can produce intensity variations which shift the apparent $\Omega(hkl)$. Since the strains, $\Delta d/d_o$, are on the order of 0.0001, a small anomalous shift nullifies the stress measurement.

2. Energy-dispersive Neutron Diffraction

To avoid the difficulties described in the previous sub-section, we have implemented and tested a mode of diffraction measurement which is "unconventional" for residual stress determination. In our measurements, we have made use of the fact that Bragg condition resonances can also be observed at fixed Ω with varying wavelength. With the scattering angle fixed, changing attenuation and texture gradients are, conceptually, less distortive. In addition, the examined volume ΔV, remains exactly the same throughout each scan.

The instrument used for energy-dispersive neutron diffraction (EDND) is shown schematically in Figure 2. Crystals of known d-spacing are placed before (monochromator) and after (analyzer) the sample; the Bragg relation is then used to select and step the wavelength incident on the sample. In principal, the analyzer crystal--which we step at the identical wavelength as the monochromator--is not needed. However, utilization of the analyzer significantly enhances instrumental resolution and produces peak profiles which are Gaussian in shape and straightforward to analyze with least-squares curve-fitting techniques.

In our system, pyrolytic graphite crystals were used for monochromator (002 plane) and analyzer (004) with a pyrolytic graphite filter between monochromator and sample. The collimation employed was 50'-20'-27'-80' from source to analyzer, with a resultant resolution $\Delta\lambda/\lambda = 0.0073$ at $\lambda = 2.69$ Å. The Cd absorbers were cut to produce rectangular apertures ranging from 1 x 30 mm^2 to 2.5 x 5 mm^2 which were used according to sample geometry and expected stress gradients.

3. Samples

A total of three different sample types were studied: a highly textured aluminum shrink-fit ring-plug specimen, a highly neutron-attenuating steel bar in a fixture for elastic bending, and two cylindrical uranium alloy samples with differing thermomechanical histories. The aluminum and steel specimens, with known, calculated stress fields, were chosen to test the technique for highly textured and highly attenuating samples, respectively. Results for these samples were in excellent agreement with theoretical predictions and demonstrated the utility of the EDND technique. Details of these measurements are described in reference 4.

For DU, two samples were studied which differed as summarized in Table II. Each specimen, the geometry of which is shown in Figure 3, was cut from the central region of 46 cm long rods.

Table II. U^{238}-0.75 wt% Ti Samples

Property/Treatment	DU1	DU2
Extrusion (10X Reduction in Area)	X	X
Solutionizing (2 hrs., 850°C, vacuum)	X	X
Furnace Cooled	X	
Quench		X
Rotary Straightened		X
Aged		X
α-U + TiU$_2$ Structure	X	
α'-U Structure		X

For convenience, we designate the furnace-cooled samples as DU1 and the rapidly-quenched sample as DU2. The stress distributions in these specific samples were not known—for the reasons detailed in the Introduction. However, based on studies of steel and other samples, it was expected that the furnace-cooled sample would be virtually stress free and the rapidly-quenched sample would have a quenched-cylinder type of distribution.

4. Stress-Strain Relations

The relation between stress and strain applicable to diffraction measurements has been presented by Evenschor and Hauk (6). With reference to Fig. 4, the $\vec{P}_i$ represent specimen axes, the $\vec{L}_i'$ represent laboratory axes, and the strain $\varepsilon_{\phi\psi}$ is measured along $\vec{L}_3'$; then

$$\varepsilon'_{\phi\psi} = (d_{\phi\psi} - d_o)/d_o$$

$$= (\epsilon_{11}\cos^2\phi + \epsilon_{12}\sin2\phi + \epsilon_{22}\sin^2\phi)\sin^2\psi \qquad (1)$$

$$+ (\epsilon_{13}\cos\phi + \epsilon_{23}\sin\phi)\sin2\psi + \epsilon_{33}\cos^2\psi$$

where $d_{\phi\psi}$ is the lattice spacing along $\vec{L}_3'$ and d_o is the unstressed lattice spacing. The stresses are related to the measured strains through

$$\epsilon_{\phi\psi}' = 1/2\ S_2(hkl)[\sigma_{11}\cos^2\phi\sin^2\psi + \sigma_{22}\sin^2_\phi\sin^2\psi$$

$$+ \sigma_{33}\cos^2\psi + \sigma_{12}\sin2\phi\sin^2\psi + \sigma_{13}\cos\phi\sin2\psi \qquad (2)$$

$$+ \sigma_{23}\sin\phi\sin2\psi] + S_1(hkl)[\sigma_{11} + \sigma_{22} + \sigma_{33}].$$

For the cylindrical samples studied, we assume that the principal axes of stress in the samples coincide with the cylindrical geometry such that no shear stresses are present in a cylindrical coordinate system. Under this assumption, we rewrite eqn. (2) for convenient ψ-values as:

$$\epsilon_{\phi,90°}' = 1/2\ S_2(hkl)[\sigma_{rr}\cos^2\phi + \sigma_{\theta\theta}\sin^2\phi] + S_1(hkl)[\sigma_{rr} + \sigma_{\theta\theta} + \sigma_{ZZ}] \qquad (3a)$$

and

$$\epsilon_{\phi,0°}' = 1/2\ S_2(hkl)[\sigma_{ZZ}] + S_1(hkl)[\sigma_{rr} + \sigma_{\theta\theta} + \sigma_{ZZ}]. \qquad (3b)$$

The $S_i(hkl)$ are diffraction elastic constants ("XEC") for the (hkl) reflection which for an elastically isotropic solid are

$$1/2\ S_2(hkl) = (1 + \mu)/E \text{ and } S_1(hkl) = -\mu/E \qquad (4)$$

where μ, E are Poisson's ratio and Young's modulus, respectively.

5. Results and Discussions

A powder diffraction pattern showed the (111), (112), and (131) peaks to be well-resolved and reasonably intense. In order to examine an (hhh) or (hoo) reflection to further minimize texture effects (7), the (111) reflection was chosen for strain measurements. Because of the spectrometer configuration and limits, a scattering angle of 72.4° was used in this case rather than 90°. To maximize intensity, the differential volumes shown schematically in Figure 3 for scattering vector $\vec{Q}$, i.e. probe direction, perpendicular to or parallel to the cylinder axis were used. Implicit in this choice is the assumption that strains in the r,θ plane vary slowly with Z for a given r,θ.

Measurements were made at points in one plane along arbitrarily chosen orthogonal axes as indicated in Figure 3. For DU1, which was expected to

be relatively stress free, a total of seven strain measurements were made
at three points in the plane. For DU2, three strain measurements were made
at each of the nine points arrayed as in Figure 3. The d(111)-spacings for
all the DU1 measurements and a portion of the DU2 measurements are shown in
Figure 5. It is clear that DU1 and DU2 exhibit very significant
differences, assumed to arise from residual stresses.

To extract three-dimensional residual stress distributions from the d-
spacing data, eqns. 3a and 3b are used. Measurements of $\epsilon'_{0°,0°}$ and $\epsilon'_{\phi,90°}$,
the latter, for example, at $\phi = 5$ and $30°$ in DU2, yields three equations
and three unknowns. However, to determine strain, d_0, the d-spacing for
the stress-free material must be known. As indicated in Figure 5, d_0(111-
DU1) is obtained from the average of the several d-spacing measurements at
different positions in the sample. A value for d_0(111-DU2) was obtained by
means of high-resolution powder neutron diffraction measurements on a
smaller sample (1 cm diam. x 1 cm long) of the same stock, with sample-
spinning to average out texture effects. Since neither calculated nor
measured XEC were available, isotropic elastic constants were used in eqn.
4, with E = 165480 MPa and μ = 0.21 (from Metals Handbook, Vol. 1, 8th Ed.,
1961, p. 1226). The resultant σ_{rr}, $\sigma_{\theta\theta}$, and σ_{zz} are shown in Figure 6 for
several points in DU2 and the axial point of DU1.

The measured stress distributions confirm the applicability of the
EDND technique to DU. As expected, DU1 is essentially stress free as
manifested by the axial-point stress shown in Figure 6, but shown more
convincingly by the d_0(111-DU1) of Figure 5. Overall, the stress
distribution obtained for DU2 is in good agreement with what would be
expected for a rapidly-quenched cylinder and with Sach's boring-out results
(2). However, clear asymmetries in stress distributions are present as
indicated by the fact that measured stresses along the two orthogonal
diameters ("X" and "Y") do not coincide (i.e., $\sigma_{rr}^X \neq \sigma_{rr}^Y$ etc.). These
asymmetries may arise from the rotary straightening of DU2; however, the
rods are, in general, found to be bent when removed from the quench bath.
This suggests that either quenching is non-uniform (possible because many
rods are "bundled") or texture in the uranium introduces anisotropy in the
stress distribution.

It is most important for future work to note that neutron diffraction
is a technique by which texture can also be measured nondestructively.
Therefore, the role of both texture and residual stresses and their
interaction can be isolated in each step of the fabrication cycle for a
single sample. It is equally important that in the measurements performed
thus far, no special effort was made to characterize "near-surface"
stresses. With the experience now in-hand, it is expected that
measurements to within 0.5 mm of the surface of cylindrical penetrator
stock may well be possible.

CONCLUSIONS

In the present work a new, nondestructive approach--energy dispersive neutron diffraction--has been conceived, developed, and tested for subsurface residual stress measurements for textured, highly-attenuating metallurgical samples. The technique has been used to characterize subsurface residual stress distributions in depleted uranium-- nondestructively and for all principal stress directions for the first time. Because of its nondestructive nature, neutron diffraction has the potential of providing information on texture and residual stresses for all phases of the fabrication process never before obtainable for depleted uranium. In view of the projected expanding use of DU, this development should have a major impact on improving performance, durability, and economics of all future DU-based armament components.

ACKNOWLEDGEMENTS

REFERENCES

1. C. P. Gazzara and R. E. Hinxman, III, "Measurement of X-ray Residual Stress in Depleted Uranium Penetrators (XM833)," AMMRC Tech. Rept., in press.

2. F. A. Simonen, K. R. Kjarmo, C. J. Morris, K. O. Nelson, E. M. Patton, L. A. Strope, and C. A. Williams, "Residual Stresses in Depleted Uranium Bars," Battelle Pacifac NW Lab Rept. AM&S No. 5, September 1982.

3. C. S. Choi, H. J. Prask, S. F. Trevino, H. A. Alperin, and C. Bechtold, "Application of Neutron Diffraction to Nondestructive Testing Problems," NBS Tech. Note 995, 1979, pp. 34-39.

4. H. J. Prask and C. S. Choi, "NDE of Residual Stress in Uranium by Means of Neutron Diffraction," J. Nucl. Matls., submitted for publication.

5. H. J. Prask and H. A. Alperin, unpublished work, 1978.

6. P. D. Evenshor and V. Hauk, Z. Metallkde. 66, 1975, 167-8.

7. H. Dölle, J. Appl. Cryst. 12, 1979, 489-501.
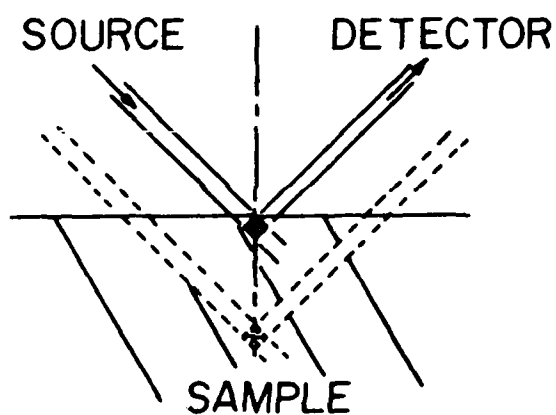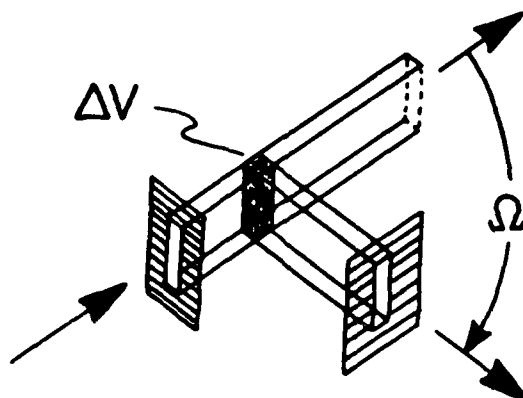
1. Upper: Perspective schematic of how absorbing masks (e.g. Cd) can be used with Bragg's Law to define the examined volume; lower: plan view of movement of examined volume through the (translated) sample.
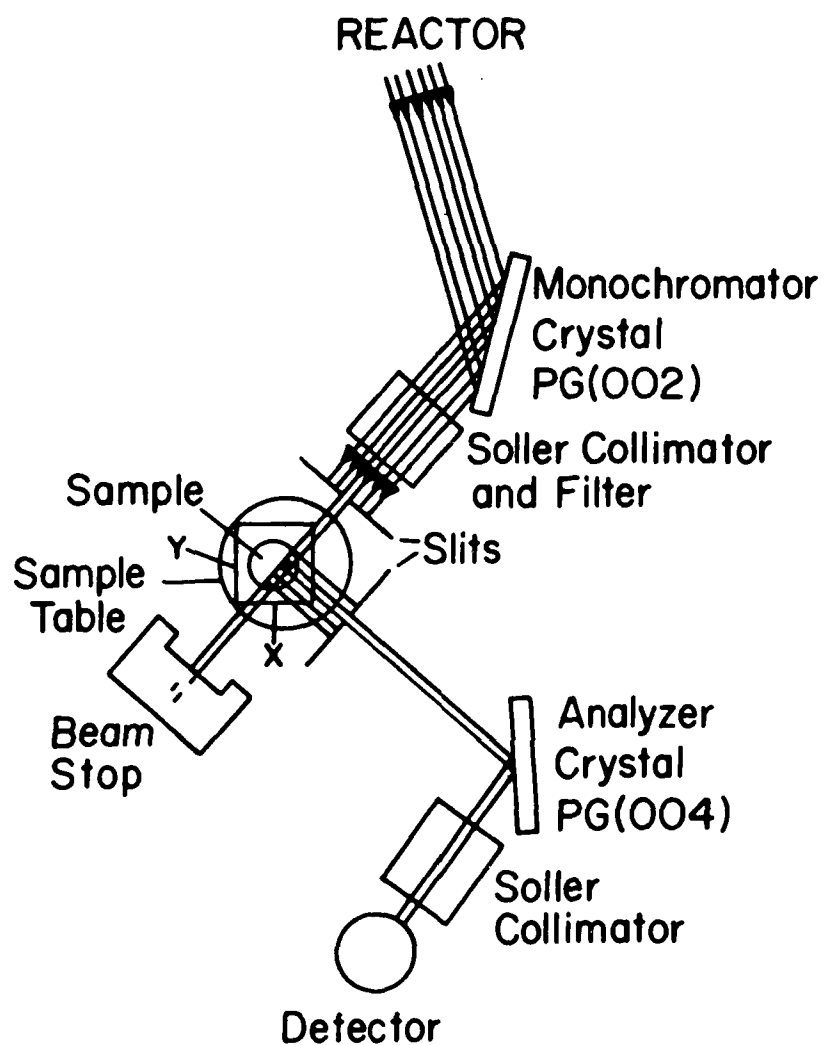
2. Schematic representation of the triple-axis neutron spec-
trometer employed for energy-dispersive diffraction.
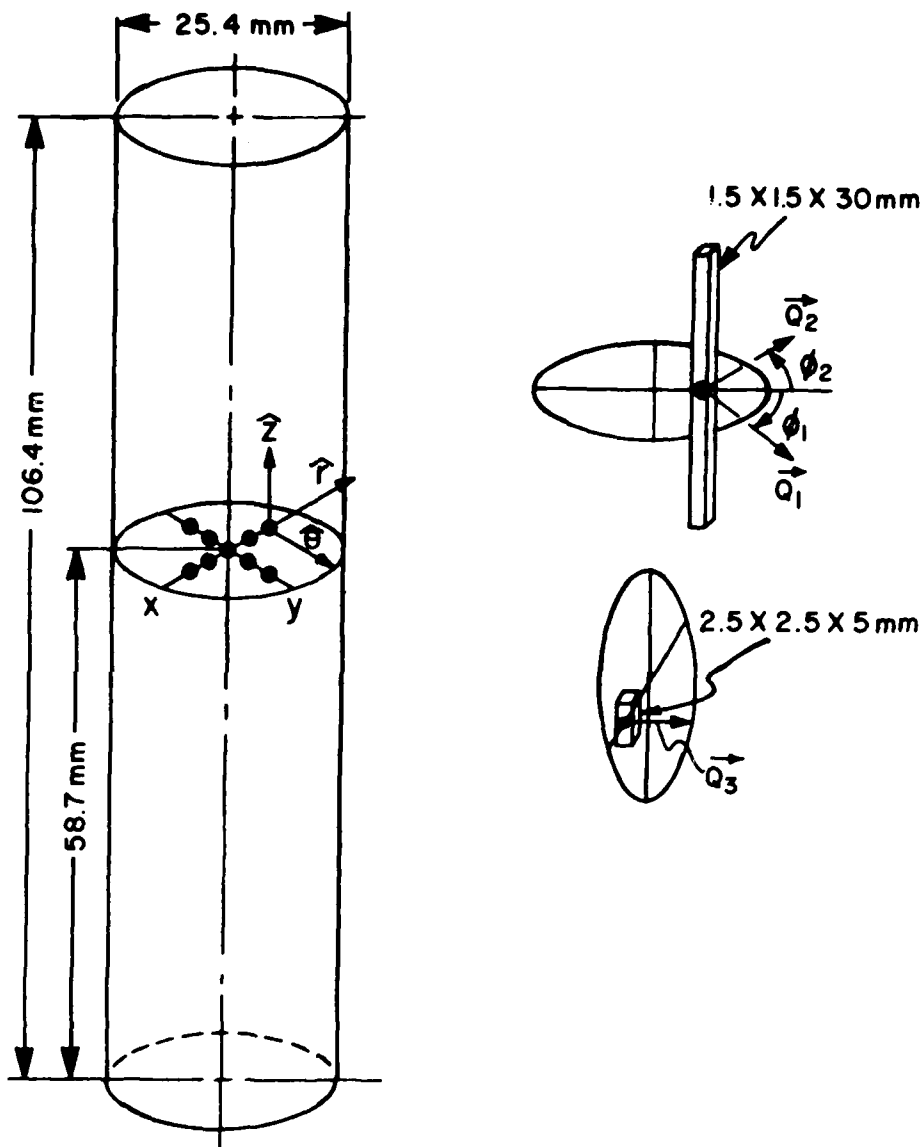
3. Overall geometry of the DU samples and points at which measurements were made. Also shown are the differential volumes examined for $\vec{Q}$s in the r-θ plane (upper right) and parallel to the cylinder axis ($\vec{Q}_3$).
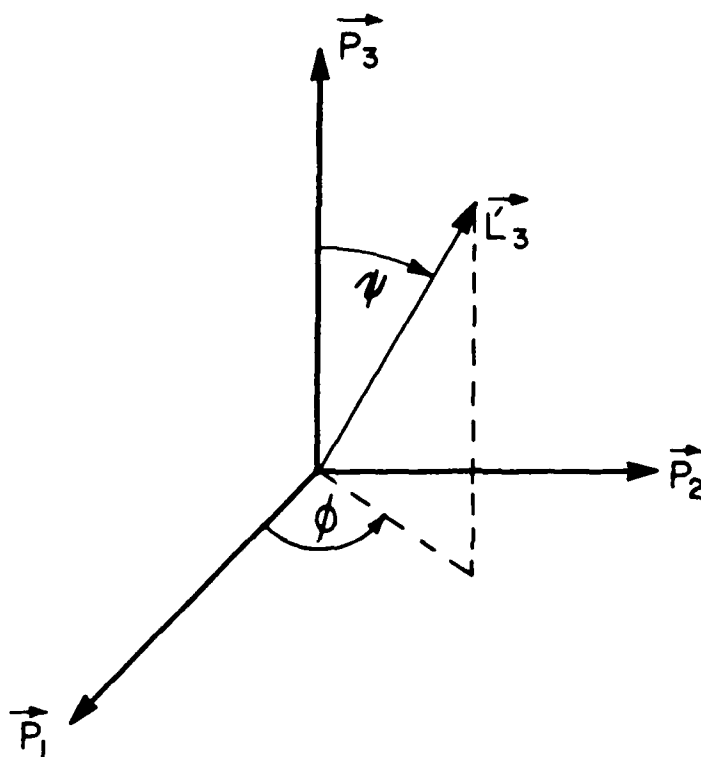
4. Specimen and laboratory coordinates at a point in the specimen. The $\vec{L}_3'$ (or probe) direction coincides with the momentum transfer, $\vec{Q}$, direction.
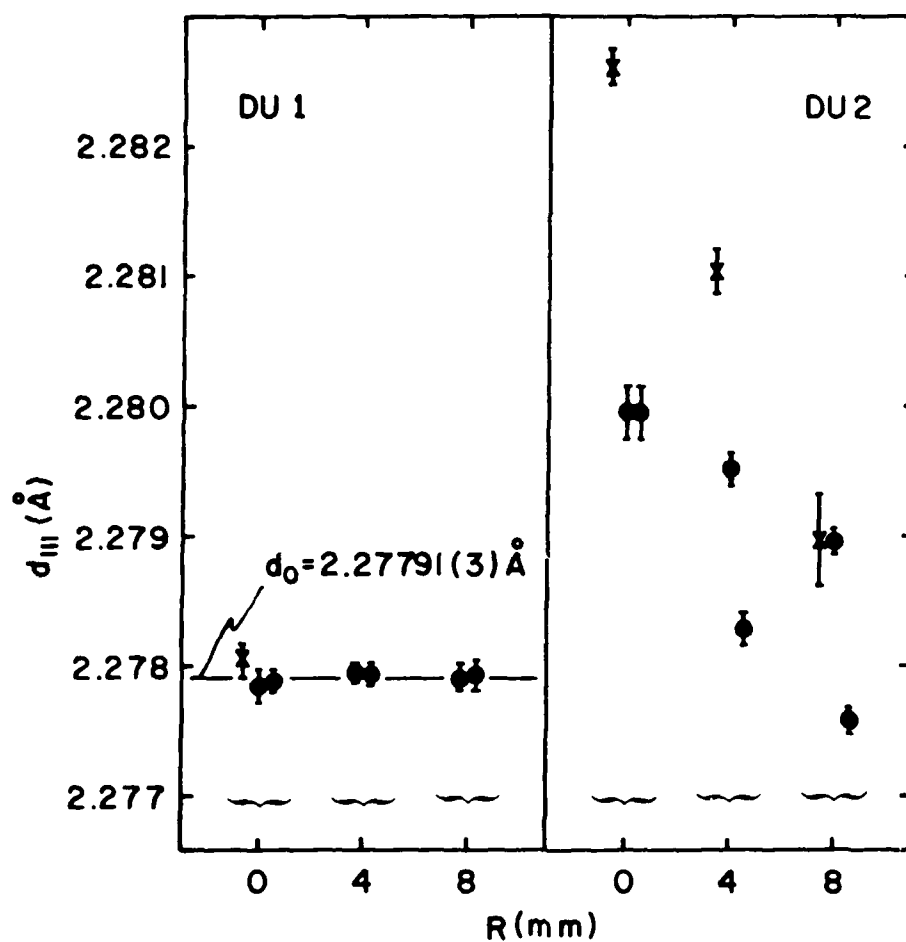
5. Selected (111) d-spacings for the furnace-cooled (DU1) and
   quenched (DU2) specimens. Measurements with $\vec{Q}||Z$ are shown
   as Xs.
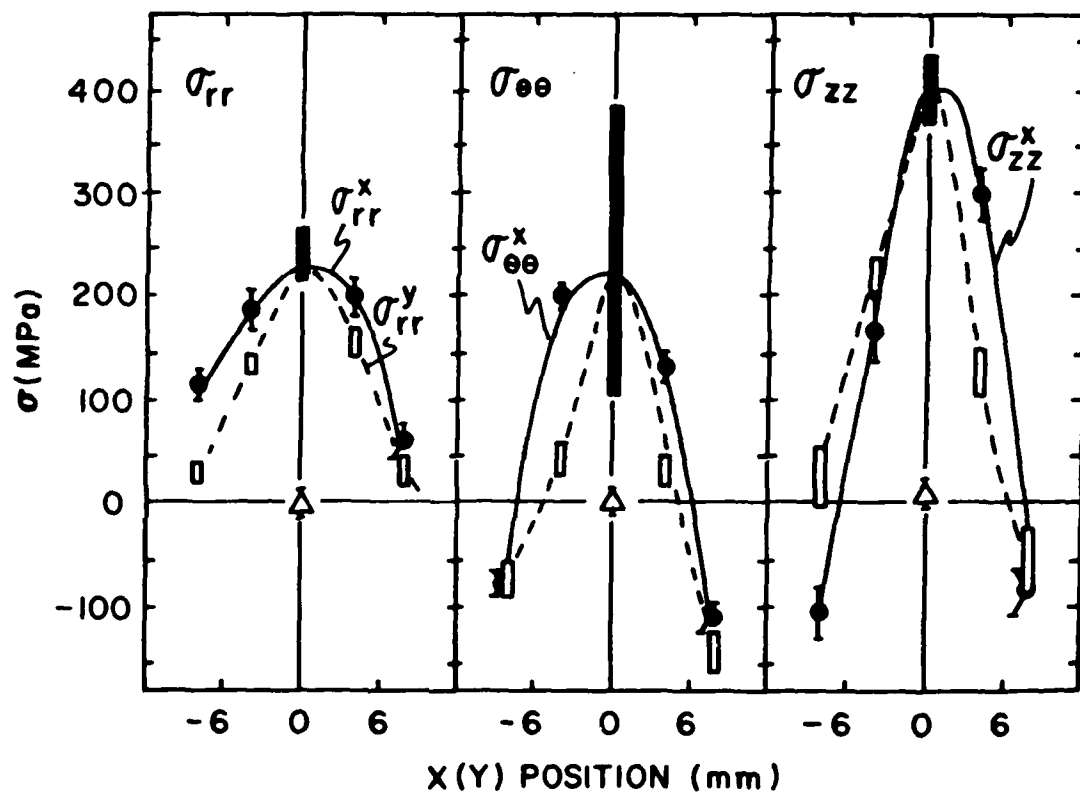
6. Residual stresses in the two depleted uranium samples [DU1
    results are shown as triangles at x(y) = 0]. The procedure
    employed is described in the text.

USE OF ARTIFICIALLY FROZEN SOIL TO HARDEN
UNDERGROUND MISSILE SITES (U)

Mr. John H. Rand
USA Cold Regions Research & Engineering Laboratory
Hanover, New Hampshire 03755
Mr. Scott E. Blouin
Applied Research Associates, Inc.
South Royalton, Vermont 05068

INTRODUCTION

The use of underground missile complexes to shelter intercontinental ballistic missiles has created many engineering problems. Among the most essential operations within the missile sites is the ability to reliably transport a missile to the surface through a tunnel called an egress shaft. It is assumed that the underground installation consists of a main horizontal tunnel deep beneath the surface. It is also assumed that vertical egress shafts would be bored using tunnel boring machines (TBMs) and that each shaft would have installed the necessary refrigeration lines to freeze the saturated soil that is backfilled into the shaft. Once the shaft is completely backfilled, then a surface refrigeration plant would freeze the saturated soil. This should reinforce the shaft to preclude damage from an attack and the frozen backfill could be used to dissipate the waste heat produced by the life support system during a post-attack period.

One design constraint on such an installation is that the complex must continue to perform its mission up to one year in a totally self-sufficient manner or buttoned-up mode. If there was an attack on the complex, all surface support systems including the refrigeration plant would be lost. This would require that either a backup system be installed in the underground complex or that the frozen backfill material be thermally stable (i.e. will not thaw) for the hypothetical one year buttoned-up period.

Frozen soil, because of its high strength and superior energy attenuation characteristics, makes an excellent backfill material to reinforce the pre-bored shafts. Analytical studies were made on the thermal characteristics of a 2000-ft vertical egress shaft that was backfilled with various materials. In addition to material considerations, other variables looked at were the properties of the surrounding rock, the freezing times

and energy consumption, and the freezing schemes as well as the thermal stability of the shaft assuming that no refrigeration was provided after one year of freezing. To support the frozen backfill egress concept, dynamic stability studies of the frozen shaft when subjected to a nuclear surface and air burst have been made.

The concept of using artificially frozen soil to harden underground installations is so attractive because the initial shafts can be bored when surface power is available, the frozen soil provides the structural stability to survive a direct attack on the installation, the shafts can remain thermally stable for the one year buttoned-up period, and the egress shaft could provide additional heat sink capabilities for the installation.

APPROACH

There are a number of variables that determine the freezing characteristics of a backfilled shaft. They include the size of refrigeration system, the refrigeration pipe layout within the shaft, the geometry of the shaft, the thermal properties of the backfill material, the thermal properties of the surrounding rock and whether or not insulation is used in the shaft.

Two piping concepts were considered for this analysis. The first is a central system that has two concentric freezing pipes along the axis of the shaft. Cold brine circulates through the inner section of the pipe and returns along the outer annulus of the pipe. The second system has 24 evenly spaced pipes running axially along the shafts' perimeter. Each set of two adjacent pipes forms a loop for the extraction of heat from the surrounding material.

Two diameters of the egress shaft were analyzed, but only the analysis of the 15 ft shaft will be described here. Because survival constraints will probably dictate that the installation be buried deeper than 2000 ft, that length was used as an arbitrary depth.

A variety of backfill materials was considered. The surrounding rock material selected was sandstone, which exhibits extreme variations in thermal properties, depending on its water content.

THERMAL CODE DEVELOPMENT

The thermal analysis used a finite element thermal code which models phase changes and incorporates nonlinear thermal properties (TEMP) (1). Details of the finite element formulation incorporated in the code are as follows.

The transient heat conduction equation is expressed as a linear differential equation using the finite element space discretization (2)

$$[C]\ \{\overset{\circ}{T}\} + [K]\ \{T\} = \{Q\}$$

where [C] is the heat capacity matrix, [K] is the conductivity matrix, {Q} is the heat flux input vector, {T} is the temperature vector and the super-posed dot denotes differentiation with respect to time.

Assuming linear variation of nodal temperature and heat flux between the time step t and t+ t, and setting the equilibrium at time t ± t, we have the following general two-time-level recurrence equation:

$$\left(\frac{1}{\Delta t} [C] + \Theta [K]\right) \{T_{t+\Delta t}\} = \left(\frac{1}{\Delta t} [C] - (1 - \Theta) [K]\right) \{T_t\}$$

$$+ (1 - \Theta) \{Q_t\} + \{Q_{t+\Delta t}\}$$

where $\Theta = 0$: Euler scheme
$\quad\quad$ 1/2 : Crank-Nicolson scheme
$\quad\quad$ 2/3 : Galerkin scheme
$\quad\quad$ 1 : Backward difference scheme

The stability and accuracy of each scheme have been discussed (2,3).

A three-time-level difference scheme has been proposed (4) and satis-factorily used in phase change problems (5,6). This difference scheme yields the following recurrence equation:

$$\left(\frac{1}{2\Delta t} [C] + \frac{1}{3} [K]\right) \{T_{t-\Delta t}\} = -\frac{1}{3} [K] \{T_t\}$$

$$+ \left(\frac{1}{2\Delta t} [C] - \frac{1}{3} [K]\right) \{T_{t-\Delta t}\} + \{Q_t\}.$$

One of the advantages of this algorithm is that it avoids iterations in non-linear heat conduction problems because equilibrium is established at the intermediate time level. The additional starting temperature re-quired can be approximated by one of the two-time-level difference schemes. Both two-time- and three-time-level difference schemes are programmed into TEMP and are available as options.

The latent heat required for the phase change of water or saturated earth materials from the unfrozen to the frozen condition can be represent-ed by a sudden jump in heat capacity over a finite temperature range cen-tered about the freezing point. To avoid the numerical difficulties associated with this large jump, a heat capacity averaging scheme based on the spatial distribution of the enthalpy gradient with respect to the temperature in an element was introduced (5). In two-dimensional problems, the average heat capacity in an element experiencing phase change is ex-pressed as follows:

$$\langle \rho C \rangle_{x,y} \simeq \frac{1}{2} \left(\frac{\partial H/\partial x}{\partial T/\partial x} + \frac{\partial H/\partial y}{\partial T/\partial y}\right)$$

where H is the enthalpy change and is defined as the integral of heat capacity as a function of temperature over the temperature range given.

Another point to be considered in artificial freezing of saturated earth material is realistic modeling of the freezing pipe. In the formulation, the heat flow transferred from the saturated earth material to the freezing pipe can be approximated by

$$Q = Ah \, (T_o - T)$$

where A = the area of surface of contact
h = the heat transfer coefficient
$T = (T + T_k)/2$
$T_o = (T_i + T_j)/2$.

The inclusion of this heat transfer mechanism requires modification of conductivity matrix and heat flux vector. The fluid temperature along the pipe can be approximated by a heat balance, considering the mass transfer of fluid.

According to trial analyses for problems involving phase change, the backward difference scheme yields the most accurate and stable results (1).

THERMAL PROPERTIES OF THE BACKFILL MATERIALS

The thermal properties of the various materials having potential use as frozen backfill in the egress tunnels, as well as a range of thermal properties of the host rock with the extremes represented by dry and saturated sandstones, are defined. Materials considered include ice, frozen silt and frozen sand. The thermal properties of ice are documented in the literature (7). Assumptions and calculations regarding the thermal properties of the other backfill materials and host rock are summarized in Table 1.

Table 1. Summary of thermal properties (n = porosity; k = thermal conductivity [Btu/ft hr °F]; C = volumetric heat capacity [Btu/ft³ °F]; λ = volumetric latent heat [Btu/ft³]).

Material	Frozen		Unfrozen		Latent heat
	k	C	k	C	λ
Ice-water	1.39	27.0	0.35	62.4	8990
Silt (n = 50%)	1.78	28.0	0.89	45.7	4500
Sand (n = 35%)	1.91	28.3	1.18	40.7	3150
Saturated sandstone (n = 15%)	2.11	28.7	1.72	34.00	1350
Dry sandstone (n = 20%)	0.51	23.8	0.51	28.8	0

RAND & BLOUIN

Semi-empirical equations describing the thermal conductivity of granular soils are presented in the literature (8,9,10). Johansen's geometric mean equation gives the best agreement with measured values for both frozen and unfrozen saturated granular soils (11).

For saturated unfrozen soils,

$$k_{sat} = k_s^{(1-n)} \, k_w^n$$

where k_s = effective solid thermal conductivity
k_w = thermal conductivity of water
n = soil porosity.

For saturated frozen soils,

$$k_{sat} = k_s^{(1-n)} \, k_i^{(n-W_u)} \, k_w^u$$

where k_i = thermal conductivity of ice
W_u = unfrozen water content as a fraction of unit soil volume

$$k_s = k_q \, k_o^{-q}$$

where k_q = thermal conductivity of quartz
k_o = thermal conductivity of solids other than quartz
q = quartz content as a fraction of the total solids content.

Thermal conductivity is a measure of the ease with which heat energy is transmitted through a material. Thermal conductivities for silt with assumed porosity, n, of 50%, sand (n = 35%), and sandstone (n = 15%) are calculated based on Johansen's equation assuming 50% quartz content (3).

The computed thermal conductivities are close to the corresponding measured values (8,9,12,13,14).

The volumetric heat capacity (C) and the volumetric latent heat (λ), are computed from the following equations:

in the unfrozen state,

$$C = \frac{\gamma_d}{\gamma_w} \, (c_s + w) \, C_w$$

in the frozen state,

$$C = \frac{\gamma_d}{\gamma_w} \, (c_s + wG_i c_i) \, C_w$$

and

$$\lambda = \gamma_d L w$$

97

where γ_d = dry bulk density (g/cm³)
γ_w = density of water (g/cm³)
c_s = specific heat of solid grains (cal/g °C)
c_i = specific heat of ice (cal/g °C)
G_i = specific gravity of ice
w = water content
C_w = volumetric heat capacity of water (cal/cm³ °C)
L = latent heat of water (cal/g).

For dry sandstone, laboratory thermal properties for a sandstone having an average porosity of 20% have been used (15).

Proper modeling of the heat transfer from the warm surrounding medium to the freezing system is one of the considerations in numerical analysis of artificial freezing. The cold brine is pumped through steel freezing pipes that are embedded in the backfill material. Boundary layers in the brine lines are created near the wall of the pipe by drag between the wall and the moving brine. The inhibiting heat transfer characteristics of the boundary layers are expressed in terms of a convective heat transfer coefficient h for the brine/pipe wall interface. This coefficient can be expressed as a function of the brine properties, the flow velocity and the pipe diameter. It is computed from the Nusselt equation from Dittus and Boelter (16):

$$h = 0.023 \frac{k}{D} \left(\frac{VD\rho}{\mu}\right)^{0.8} \left(\frac{c\mu}{k}\right)^{0.4}$$

where D = diameter of the pipe (ft)
ρ = mass density (lb/ft³)
V = velocity (ft/hr)
μ = absolute viscosity (lb/ft hr)
k = thermal conductivity (Btu/hr ft °F)
c = specific heat (Btu/lb °F).

In addition to the convective heat transfer coefficient for the boundary layer in the brine, the conductivity of the steel freezing pipe also influences the heat transfer to the brine. Pipe conductivity is included in the overall heat transfer coefficient U for a given freezing pipe expressed as:

$$U = \frac{1}{\dfrac{r_0}{hr_1} + \dfrac{r_0 \log_e \left(\dfrac{r_2}{r_1}\right)}{k}}$$

where r_1 = inner radius of pipe
r_2 = outer radius of pipe
$r_0 = (r_1 + r_2)/2$
k = thermal conductivity of steel pipe (assumed 26 Btu/hr ft °F).

THERMAL ANALYSIS RESULTS

The time required to completely freeze the backfill using the central pipe system is strongly dependent on heat flow from the surrounding rock. With perfect insulation surrounding the backfill, freezing time is about 150 days in dry sandstone and about 440 days in wet sandstone.

The power required to freeze the saturated sand backfill using the central pipe system in the insulated case is 0.9×10^6 Btu per lineal foot of egress shaft (260 kWh/ft). Likewise, 1.3×10^6 Btu/ft (390 kWh/ft) is required to freeze the uninsulated backfill in dry sandstone, and 3.6×10^6 Btu/ft (1050 kWh/ft) is required to freeze the saturated sand backfill in the saturated sandstone. Because of the continuous heat flow into the un-insulated frozen backfill from the surrounding rock, additional power is required to maintain the uninsulated backfill in the frozen state. The maintenance power decreases gradually with time as the surrounding rock is cooled. At 400 days, maintenance power needed in the wet sandstone is 0.11 hp per lineal foot of egress shaft, compared to 0.07 hp/ft in the dry sandstone.

The peripheral freezing pipe system, in the uninsulated case, freezes the backfill and surrounding rock simultaneously. The freezing front propagation in the backfill is independent of the thermal properties of the surrounding rock. Freezing time for the peripheral system in the saturated sand backfill is about 20 days, compared to about 28 days in the saturated silt and about 67 days in the water. The primary reason for the difference is dependence of the latent heat of freezing on the water content. The saturated silt contains more water per unit volume than the saturated sand. In the uninsulated case, the wet sandstone takes 60 days to freeze to a distance of 7.5 ft (1 radius) from the shaft wall. The corresponding time to reach a temperature of 32°F at that range in the dry sandstone is 130 days. The dry sandstone cools less rapidly because of its lower conductivity.

Because of the greater surface area of the peripheral system, and because both the backfill and surrounding rock are being cooled simultaneously, its power requirements are significantly higher. With the assumed -30°F brine temperature, the backfill freezes rapidly; and the initial power requirements decrease steeply to much lower maintenance levels within 50 days.

For the purposes of using efficiently designed refrigeration systems, the initial freezing times would be lengthened to permit use of a smaller capacity refrigeration system. For instance, a 1000 hp refrigeration system would be sufficient to freeze and maintain an uninsulated 2000 ft egress shaft in saturated sandstone. However, initial freezing of the sand backfill would take somewhat longer than the 20 days. A system of about 300 hp would be sufficient to freeze and maintain the sand backfill in the dry sandstone. The maintenance power decreases quicker in the case of the peripheral freezing pipe system, so that after several years, maintenance power requirements will approach those for the central pipe system.

The peak power required to freeze and maintain the central pipe system backfilled with saturated sand would range from about 300 to 400 hp, depending on the initial freezing criteria. This requirement is based on the energy actually needed to freeze the backfill and does not take into account inefficiencies in the refrigeration system.

The thermal stability of frozen egress shafts was analyzed. This assumed that the refrigeration plant was destroyed by an attack after one year of operation with the initial temperature conditions based on the calculated temperature profiles at that time. The thawing front reaches the uninsulated backfill about 12 months following the attack in the case of dry sandstone and about 18 months following the attack in the case of the wet sandstone. Complete natural thawing of the backfill takes 29 months in the wet sandstone and 33 months in the dry sandstone. Nearly the full reinforcement and strength characteristics of the frozen backfill egress system will be maintained for a least a year following an attack.

GAINING ACCESS THROUGH THE FROZEN BACKFILL MATERIAL

Several methods to gain access through the frozen backfilled egress shafts have been considered. The method most discussed is mechanically excavating the frozen backfill using the TBMs. The installation, as well as the egress shaft, will initially have to be bored using such a device. The equipment, which is state of the art equipment and commercially available, will consume a substantial amount of power to operate during the egress phase. The disadvantage of using this type of equipment is that large power requirement.

Data collected from TBMs indicate that the power consumed is roughly proportional to the tunnel face area and the power density ranges from 0.02 to 0.06 hp/in^2. The tunnel face area is 25,400 in.^{2}. Assuming an average power density of 0.04, then the power required to excavate would be roughly 1000 hp (750 kW). Estimates of 600 kW for 150 hours have been discussed at various deep-basing briefings. In order to accomplish this, assuming a 2000-ft tunnel, the TBM would have to penetrate at a rate of at least 2.6 in./min. This is extremely fast and it is doubtful that the system could maintain that performance.

Aside from the large power consumption, another problem associated with mechanically excavating is the muck removal and storage. Muck is the material that has been excavated from the tunnel. Typically, when any material is excavated it expands to about 1.8 times its original volume. In the vertical egress shaft configuration, the muck could be deposited in a prebored storage area directly beneath the egress shaft.

The second method to gain access through the frozen backfill is to use the egress tunnel as a heat sink. Existing reports are available that present the criteria, engineering information and estimation procedures for the disposal of waste heat associated with the generation of power required to supply the needs of a hardened underground defense installation (18).

Because firm data on the anticipated waste heat loads and other design constraints are not available, parametric curves on numerical models have been generated to determine various heat sink behaviors once waste heat load parameters are specified (20). From the results of the thermal analysis it is apparent that even after one year of the refrigeration system being shut off, the egress shaft remains frozen. Particular problems exist in using either the peripheral freeze pipe system or the central system as a waste heat exchanger. A secondary system might be required for the waste heat to be effectively transferred to the backfill material. A study conducted by Foster-Miller Associates (21) in 1975 looked at the ice-water heat sink concept and assessed the available options for flow paths of coolant water within the heat sink. They concluded that such ice-water heat sinks were feasible.

Use of the egress shaft as a heat sink eliminates some of the need for dedicated heat sink tunnels, which are labor-intensive and costly. Such a system would probably not be used to excavate the shaft in an expedient manner. The backfill material in the egress shaft and the design constraints on the shaft could be such that during a period of time, in the buttoned up phase, the bottom of the egress shaft backfill could be thawed by the waste heat. When it came time to dig out for the egress phase, much of the material would already be removed.

The final and most elaborate concept for a rapid egress system is to develop a system whereby heat is introduced to the interface of the shaft wall and the backfilled material. By using the peripheral freezing pipe to thaw the circumference of the egress shaft, a plug or cylinder would result. This cylinder, containing a major portion of the frozen material, would drop or be lowered into the sump tunnel located directly below the egress shaft. This method consumes less energy and could be the fastest method available to gain access to the surface. This scheme would require the removal of the frozen cylinder in sections. Vertical shafts would not be as attractive as inclined shafts because of the extreme amount of weight of each cylinder. To drop a cylinder of frozen material 200 ft long and 14 ft in diameter through the egress shaft into a matched sump tunnel in itself would be a major event. Special systems to control the descent of these plugs would be necessary.

DYNAMIC ANALYSIS

Two studies have been completed investigating the dynamic response of a frozen backfilled egress shaft and the surrounding rock to a high yield nuclear surface burst and air blast loading from a near miss surface burst (1,18). The analysis of the response of the frozen backfilled shafts to the nuclear surface burst and air blast led to the following conclusions:

1. Frozen sand backfill appears to maintain the structural integrity of the shafts under high dynamic stress of up to 3 to 6 kbars. Rockbolt reinforced structures without backfill would generally be expected to collapse at 1 kbar or less.

2. Maximum circumferential compressive strain from the airblast in-
duced loading is comparable to that from the airblast loading in the hard
rock. Maximum strain from the direct induced loading in the soft rock is
2.2% and is largely attributable to temporary pressure melting and soften-
ing of the frozen backfill. The high peak strain in the soft rock is
close to the failure strain for a rockbolt reinforced opening in this
material.

3. The presence of a backfilled shaft causes very little perturbance
of the stresses and ground motions in the rock. This lack of perturbation
indicates that the shaft will have little or no mitigating influence on
stresses and motions reaching deep based facilities located below or near
the base of the shaft.

CONCLUSIONS

As a result of this study the following conclusions are drawn:

Power requirement to freeze and maintain a 2000-ft uninsulated egress
shaft, backfilled with saturated sand, is in the range of 500 hp.

The backfill material can be frozen in approximately 1 month's time,
depending upon the physical size, arrangement of the freezing system, and
the thermal properties of the backfill and surrounding rock.

Insulating the egress shaft reduces the freezing time, lowers the
power requirements, and increases the thermal stability of the shaft dur-
ing the post attack period. The major problem with insulating the shaft
is that it will result in a major degradation in the survivability.

Uninsulated shafts are capable of remaining frozen for more than a
year, assuming the shaft has undergone at least one year of freezing.

Many options remain possible for egress plug removal, including outer
periphery thawing, heat sink capabilities and mechanical excavation.

RECOMMENDATIONS

In order to conduct this preliminary study many assumptions and over
simplifications have been made. From this, however, the frozen backfilled
egress shaft looks feasible.

Specific problems, such as how the backfill material is to be placed
in the shaft, remain unanswered. There are many of these questions that
will require separate studies. From a practical stand-point it is doubt-
ful that vertical 2000-ft shafts can be bored; maximum slopes of 40° might
be obtainable. For simplification of the analysis, however, the vertical
shaft was chosen.

Because of the insufficiencies in the study, it is recommended that
supplemental studies be conducted to increase confidence in the analysis.

LITERATURE CITED.

1. Blouin, S.E., and Kim K.J., Use of Frozen Backfill as a Combination Egress/Heat Sink System for Deep Basing, Applied Research Associates, Inc., South Royalton, Vermont, October 1983.

2. Zienkiewicz, O.C., The Finite Element Method in Engineering Science, McGraw-Hill, London, 1971.

3. Donea, J., "On the Accuracy of Finite Element Solutions to the Transient Heat-Conduction Equation," International Journal of Numerical Methods for Engineering, Vol. 8, pp. 103-110, 1974.

4. Lees, M., "A Linear Three-Level Difference Scheme for Quasilinear Parabolic Equations," Maths. Comp., Vol. 20, pp. 516-522, 1966.

5. Comini, G., Del Guidice, S., Lewis, R.W., and Zienkiewicz, O.C., "Finite Element Solution of Non-Linear Heat Conduction Problems with Special Reference of Phase Change," International Journal of Numerical Methods for Engineering, Vol. 8, pp. 613-624, 1974.

6. Hsu, T.R., and Pizey, G., "On the Prediction of Fusion Rate of Ice by Finite Element Analysis," Journal of Heat Transfer, Vol. 103, pp. 727-732, November, 1981.

7. Yen, Y.C., Review of Thermal Properties of Snow, Ice and Sea Ice," CRREL Report 81-10, 1981.

8. Johansen, O., "Thermal Conductivity of Soils," Ph.D. Thesis, Trondheim, Norway, 1975, (CRREL Draft Translation 637, 1977) ADA044002.

9. Kersten, M.S., Laboratory Research for the Determination of the Thermal Properties of Soils, ACFEL Technical Report 23, AD 712516, 1949.

10. DeVries, D.A., "The Thermal Conductivity of Soil," Translated by Building Research Station (Library Communication No. 759), England.

11. Farouki, O.T., Thermal Properties of Soils, CRREL Monograph 81-1, 1981.

12. Wolfe, L.M., and Theim, J.O., "Physical and Thermal Properties of Frozen Soil and Ice," Journal of the Society of Petroleum Engineers, pp. 67-72, March 1964.

13. Andersland, O.B., and Anderson, D.M., Geotechnical Engineering for Cold Regions, McGraw-Hill, New York, 1978.

14.   Clark, S.P., Handbook of Physical Constants, p. 471, 1966.

15.   Somerton, W.H., "Some Thermal Characteristics of Porous Rocks," Journal of Petroleum Technology, Technical Note 2008, May, pp. 61-64, 1958.

16.   Dittus, F.W., and Boelter, L.M.K., University of California Pub. Eng., 2, p. 443, 1930.

17.   Mellor, M., Mechanics of Cutting and Boring Part 7: Dynamics and Energetics of Axial Rotating Machines, CRREL Report 81-26, 1981.

18.   Stubstad, J.M., Quinn, W.F., Greenburg, M., Best, W.C., and Botros, M.M., Design Procedures for Underground Heat Sink Systems, CRREL Special Report 79-8, 1979.

19.   Kim, K.J., and Blouin, S.E., Response of Frozen Backfilled Tunnels to High Intensity Airblast Loadings, Applied Research Associates, Inc., South Royalton, Vermont, July 1983.

20.   Stubstad, J.M., Ashton, G., Yen, Y.C., Lunardini, V., Rand, J.H., and Blouin, S.E., Heat Sink Study in Support of Deep Basing for MX, CRREL Interim Report, 1982 (not published).

21.   Grande, E., Analysis and Conceptual Design of Practical Ice-Water Heat Sinks, CRREL Special Report 221, 1975.

A GENERAL SYNTHESIS OF METALLOTETRABENZPORPHYRINS (U)

*DAVID E. REMY, DR., JOHN T. STAPLER, MR.,
AARON L. BLUHM, DR., AND FRANK H. BISSETT, DR.

U.S. ARMY NATICK RESEARCH AND DEVELOPMENT CENTER
NATICK, MA 01760

INTRODUCTION

Our interest in the metallotetrabenzporphyrins stems from the fact
that they share a similar chromophore with some of the biologically
derived porphyrins such as chlorophyll. Therefore, it was felt they might
also mimic the emission characteristics of a green foliated background and
be useful in the design of a new camouflage concept. With this object in
mind we initiated a research project to study the synthesis and properties
of these compounds. During the course of this work we discovered a new
synthesis for the metallotetrabenzporphyrin system which appears to be
completely general in scope for this series of compounds and provides
exceptionally high yields (1). This report discusses this synthetic
procedure in detail, describes some of the properties of the new
compounds, and examines the potential of these structures as camouflage
dyes.

DISCUSSION

Porphyrins, phthalocyanines, and metallotetrabenzporphyrins all con-
tain pyrrole as a common structural unit. In these structures four of the
pyrrole units are linked by either carbon atoms or, in the case of
phthalocyanine, by nitrogen to form the macrocyclic ring system. The
positions occupied by these linking atoms are conventionally labeled the
meso- positions. In metallotetrabenzporphyrin and phthalocyanine
molecules the pyrrole rings form part of a larger substructure, isoindole.
All of these macrocycles can exist as their metal-free derivatives in
which the metal atom is replaced by two hydrogen atoms.

Since these compounds share a strong structural resemblance they
also share certain properties in common. The phthalocyanines and the
tetrabenzporphyrins have outstanding thermal stability. They can
withstand temperatures in excess of 500°C without degradation. Zinc

105
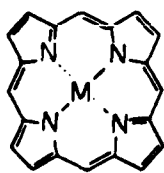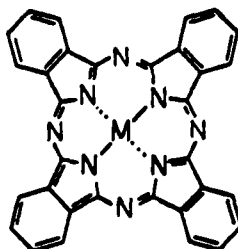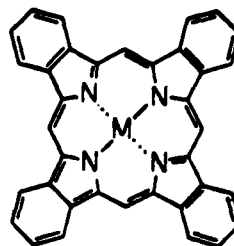
PORPHYRINS   PHTHALOCYANINES   TETRABENZPORPHYRINS

tetrabenzyporphyrin has been purified by carbonizing the impurities at
high temperature and then extracting the unaltered product from the char.
All three classes of compounds are intensely colored. The phthalocyanines
have been widely used as commercial pigments. The absorption spectra of
the porphyrins and tetrabenzporphyrins exhibit strong narrow Soret bands
around 400 nm. These bands which have exceptionally high molar
absorptivities are characteristic of the carbon linked macrocycle and are
not present, to the same extent, in the phthalocyanines.

Partly due to their biological significance, the porphyrins have
received an enormous amount of attention; and work in this important area
of chemistry continues at a rapid pace. This progress has not been
achieved by the related metallotetrabenzporphyrins, in spite of their
potential in both a theoretical and practical sense. A major reason for
the slow development of this area has been the lack of a versatile
synthetic route for these compounds. Early work was limited almost
exclusively to synthesis of the parent structure (2). Routes to substituted
tetrabenzporphyrins have been few and have generally been restricted to
the preparation of only a single derivative (3). Recently, this situation
has begun to change and derivatives are being prepared by procedures that
have some generality (4). However, these newer syntheses also have
limitations, and there still exists a pressing need for a convenient
synthetic methodology of wide applicability.

Natick's strategy for a general synthetic route relies on two main
considerations. First, inspection of the tetrabenzporphyrin structure
reveals that isoindole would be an ideal monomer for formation of the
macrocycle. This concept was previously noted but no attempt was made to
make use of this approach (5). Presumably, the previous absence of
preparative methods for isoindole and the long-held notion that
it was too reactive to be isolated or used as a synthetic precursor were
reasons why this route was not tried. Past problems with the synthesis

106

and isolation of isoindole have now been overcome and, in spite of its reactivity, it can be prepared conveniently and used as a reactant (6).

The second consideration is based on the fact that the well-studied Rothemund synthesis has been used effectively for many years for the preparation of simple porphyrins and symmetric porphyrin derivatives (7). The synthesis relies on reaction of pyrrole with an aldehyde in the presence of an acidic catalyst. A more recent modification of this classic route uses metal salts to assist formation of the macrocyclic ring and increases yields (7). This improved version of the Rothemund synthesis with the substitution of isoindole for pyrrole has been adapted by Natick to provide the first truly general synthesis of the metallotetrabenzporphyrins.

The reaction procedure can be illustrated by the preparation of zinc tetrabenzporphyrin as shown in SCHEME I where the R group on the aldehyde is hydrogen.
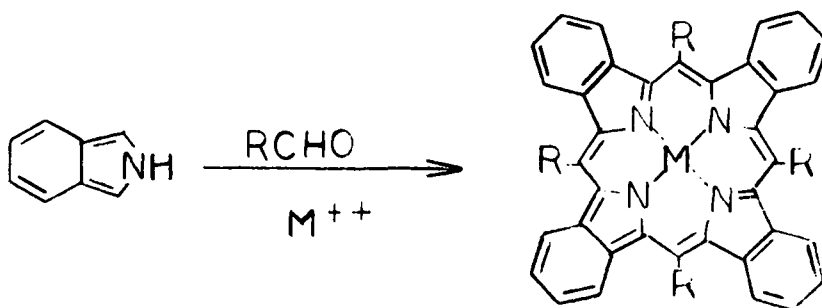
SCHEME I

In the procedure a cold solution of isoindole in degassed ether was mixed with excess zinc acetate in an inert atmosphere. Monomeric formaldehyde, prepared from paraformaldehyde, was led into the cold reaction solution with a stream of nitrogen. The temperature was raised rapidly to drive off the solvent and leave a deep purple residue, which was further heated to $375^{\circ}C$ for 20 minutes. The reaction residue was exposed to air, extracted with pyridine, and chromatographed on silica to yield purified product.

The isoindole reactant for this procedure has been prepared by two routes with equal success. One route relies on flash vacuum thermolysis of 1,2,3,4-tetrahydrohaphthalen-1, 4-imine. This route allows isolation of solid isoindole in a liquid nitrogen-cooled trap in nearly quantitative yield (9). The second route relies on the elimination of p-toluene-sulfinate from 2-(p-tolylsulfonyl)dihydroisoindole and provides an 85% yield of isoindole in ether solution (10). Overall yields for the tetra-benzporphyrin syntheses are based on the assumption of 100% formation of isoindole in the flash vacuum route and 85% in the elimination procedure.

Substitution of various aldehydes and isoindoles in the general reaction procedure has provided the compounds described in TABLE 1. Inspection of this table shows that the scope of this reaction is indeed broad and the synthesis can be used to prepare metal-free as well as metallated compounds, derivatives with meso-substituents, and derivatives with modified benzene rings. With few exceptions the yields are generally high, even with those aldehydes which have not been very successful in the classical Rothemund synthesis of simple porphyrins.

When isoindole was reacted with formaldehyde or benzaldehyde in the absence of metal, the free base derivatives were formed (Reactions 2 and 6) in 4% and 29% yields, respectively. The higher yield for the meso-phenyl derivative in this metal-free reaction can be attributed to the fact that the more bulky phenyl substituent causes the reactant to conform to a geometry which favors cyclization. When the meso-substituent is hydrogen, there is little compulsion to assume a cyclic geometry; and linear polymerization reactions can predominate, resulting in decreased yields. When metal ions are present, prior coordination of the isoindole units to the metal helps to set up the optimum geometry for cyclization; and the effects of bulky meso-substituents are less important. This theory is borne out by comparison of yields for the meso-hydrogen (53%) and meso-phenyl (50%) derivatives when excess zinc was present.

Since the participation of metal ions is clearly an important factor, the reaction of isoindole with formaldehyde was expected to be greatly assisted by the presence of nickel or copper acetates due to the tendency of these ions to adopt a square planar configuration, which would enable them to act as even better template catalysts than zinc in the formation of the porphyrin ring. Instead, only trace amounts of the expected products were obtained. This result is especially surprising in view of the relatively high yields obtained with these compounds in related syntheses. Obviously, the role of these metal salts needs to be further clarified.

A sulfonated derivative of zinc tetrabenzporphyrin was prepared by reaction with fuming sulfuric acid at room temperature. The reaction mixture was partially neutralized with aqueous base, and the acid form of the sulfonate derivative was obtained by centrifugation. Elemental analysis indicates that there is approximately one sulfonic acid group per molecule of tetrabenzporphyrin. The position of the sulfonic acid group

in the molecule was not determined, but it is presumed to be on the outer benzo- portion of the structure.

To examine the dyeing behavior and the resultant spectral reflectance properties of the tetrabenzporphyrins on a textile substrate, the sulfonic acid derivative was converted to its sodium salt by treatment with an equivalent amount of base. It was then used in a conventional dyeing process to dye a sample of NYCO fabric (cotton/nylon). NYCO is the fabric used in the camouflage battle dress uniform (BDU). The spectral reflectance curve of this dyed sample is shown in Figure 1. The curves for the green portion of the BDU camouflage pattern and for natural green vegetation are also shown in Figure 1.
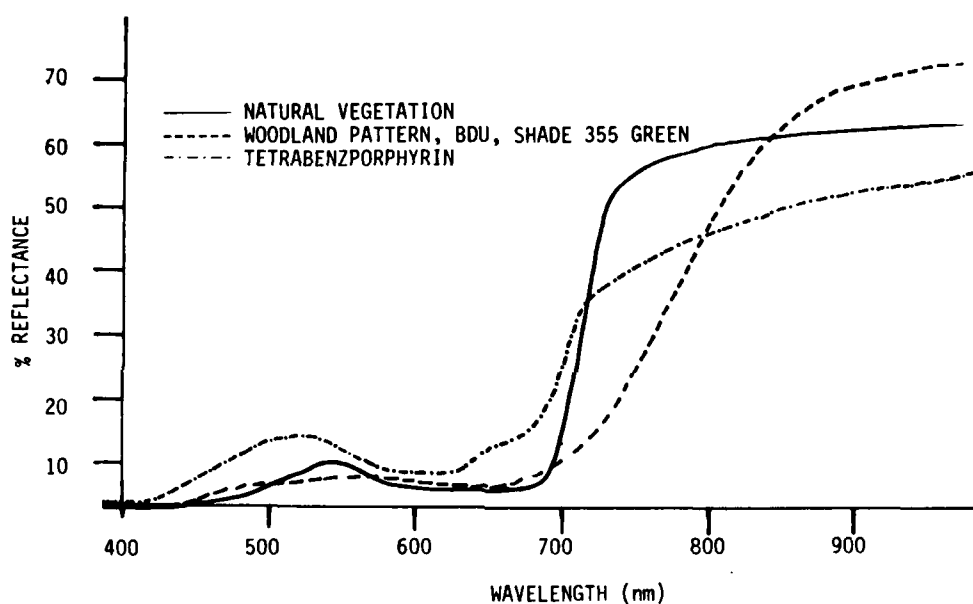
FIGURE 1

The results show that this preliminary trial with the tetrabenz-porphyrin dye compares favorably with the reflectance spectrum for natural vegetation both in the visible and in the near-infrared. The standard camouflage green pattern is an optimized mixture of several dyes and, as can be seen, it provides an average approximation of the reflectance spectrum of the vegetative background.

109

CONCLUSION

The scope of a new reaction to prepare metallotetrabenzporphyrins has been explored and many derivatives of this system have been prepared for the first time. Various metal salts have been tried in the synthetic procedure and the effects of these metal ions on the ring-closure reaction have been investigated. Zinc tetrabenzporphyrin has been converted to its sulfonated derivative and has been dyed onto a military fabric. The spectral reflectance curves for this dyed sample, the green BDU camouflage pattern and natural vegetative background have been compared.

ACKNOWLEDGEMENTS

TABLE 1

Reaction*	meso-Substituent	Metal	Yield
1	Hydrogen	Zn	53%
2	Hydrogen	2H	4%
3	Hydrogen	Ni	2%
4	Hydrogen	Cu	< 1%
5	Phenyl	Zn	50%
6	Phenyl	2H	28%
7	p-Dimethylamino	Zn	83%
8	p-Dimethylamino	2H	25%
9	p-Methoxyphenyl	Zn	25%
10	m-Nitrophenyl	Zn	< 1%
11	o-Hydroxyphenyl	Zn	< 1%
12	p-Hydroxyphenyl	Zn	2%
13	1-Naphthyl	Zn	53%
14	2-Naphthyl	Zn	78%
15	9-Phenanthryl	Zn	26%
16	9-Anthracenyl	Zn	6%
17	n-Butyl	Zn	17%
18	Methyl	Zn	42%
19	Hydrogen	Zn	57%
20	Phenyl	Zn	15%
21	Hydrogen	Zn	< 1%

*Isoindole was used as a reactant for reactions 1 to 18,
1,2,3,4-tetrafluoroisoindole for reactions 19 and 20, and
1,2,3,4-tetrabromoisoindole for reaction 21.

REFERENCES

1. D.E. Remy, Tetrahedron Lett., 1451 (1983).

2. J.H. Heilberger, Justus Liebigs Ann. Chem., 529, 205 (1937);
J.H. Heilberger and A. von Rebay, ibid, 531, 279 (1937); J.H. Heilberger,
A. von Rebay, and D.B. Hever; ibid, 533, 197 (1938); P.A. Barrett, R.P.
Linstead, F.G. Rundal, and G.A.P. Tuey, J. Chem. Soc., 1079 (1940).

3. A. Vogler and H. Kunkely, Angew. Chem. Int. Ed. Engl., 17, 760 (1978);
C.O. Bender, R. Bonnett and R.G. Smith, J. Chem. Soc., Perkin Trans. 1,
771 (1972); D. Dolphin, J.R. Sams and T.B. Tsin, in "Inorganic Syntheses",
D.H. Busch, Ed., J. Wiley and Sons, Inc., New York, NY, 1980, Vol. XX,
p155.

4. V.N. Kopranenkov, S.N. Dashkevich and E.A. Luk'yanets,
Zh. Obshch. Khim., 51, 2513 (1981); Chem. Abstr., 96, 96523d (1982).

5. L. Edwards, M. Gouterman, and C.B. Rose, J. Am. Chem. Soc., 98, 7638
(1976); R.P. Linstead and E.G. Noble, J. Chem. Soc., 933 (1937).

6. R. Bonnett and S.A. North, in "Advances in Heterocyclic Chemistry",
A.R. Katritzky and A.J. Boulton, Ed., Academic Press, New York, NY, 1981,
Vol. 29, p341.

7. P. Rothemund, J. Am. Chem. Soc., 57, 2010 (1935); ibid, 58, 625
(1936); ibid, 61, 2921 (1939).

8. R.H. Ball, G.D. Dorough, and M. Calvin, J. Am. Chem. Soc., 68, 2278
(1946).

9. J. Bornstein, D.E. Remy, and J.E. Shields, J. Chem. Soc., Chem. Comm.,
1149 (1972).

10. R. Kreher and J. Seubert, Z. Naturforsch, 20b, 75 (1965).

SIMULATING ADIABATIC SHEAR

BARBARA E. RINGERS, MRS.
US Army AMCCOM, ARDC/Ballistic Research Laboratory
Terminal Ballistics Division
Aberdeen Proving Ground, MD 21005

I. INTRODUCTION

Plugging due to adiabatic shear is a prevalent mode of target failure in both rolled homogeneous armor targets and high hard armor targets. It is therefore of considerable terminal ballistic interest to be able to accurately model this phenomenon. This paper presents techniques and criteria developed to handle simulations of target plugging failure due to adiabatic shear and compares experiment and simulations for two impacts involving titanium alloy targets.

II. SLIDING SURFACE TECHNIQUES

A Lagrangian, dynamic, impact code, EPIC-2 (1), was utilized as the basis for this effort. Major modifications had been developed and implemented by the author to provide the mechanics necessary to handle deep penetration/perforation of targets (2). Splitting between target elements (triangles in EPIC-2) was utilized to simulate shearing failure (Figure 1) and total element failure was enabled to simulate the expected accompanying erosion in both projectile and target. An automatic dynamic relocation of sliding surfaces and the automatic addition of new sliding surfaces, when appropriate, enabled both of these target failure techniques to carry a calculation to completion, i.e. through target perforation, if applicable. The beauty of these sliding surface techniques is that they provide the mechanics by which criteria may be tested for their capabilities in handling specific modes of target failure. These techniques were first successfully applied in conjunction with an equivalent strain criterion to handle the modeling of plugging failure due to high strains (3).

113

SPLIT WILL BE FURTHERED FROM "NEXT" NODE
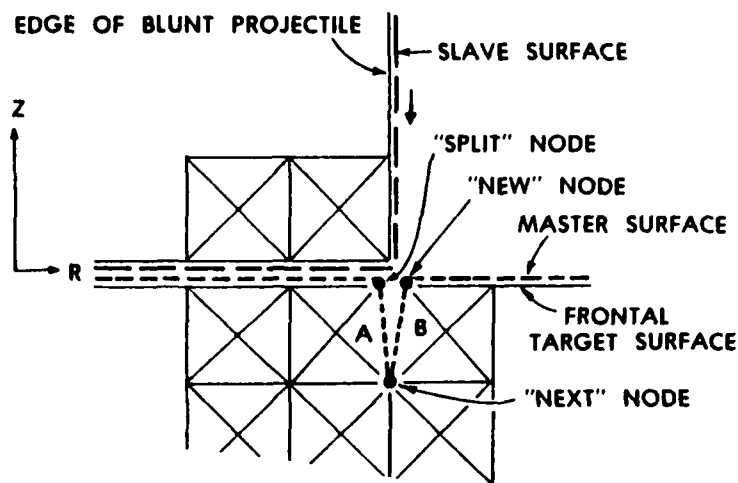
Figure 1. Splitting Between Elements

TI 318 TARGET

NO SIGN OF SHEAR BAND
BY 10 μs

TI 125 TARGET

INCIPIENT SHEAR BAND
BY 10 μs
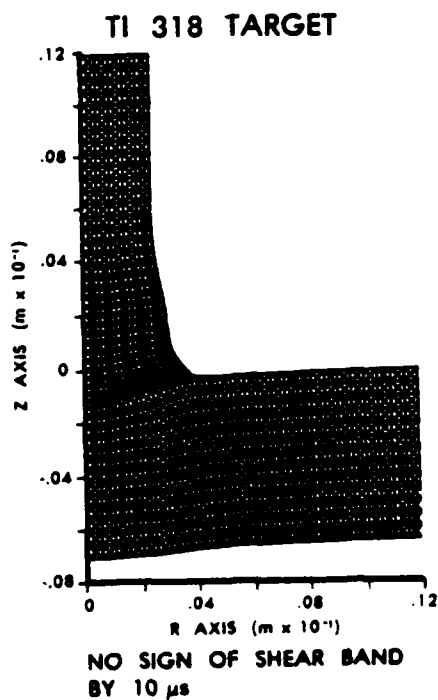
Figure 2. Impact of Steel Cylinder/
Ti 318 Target @ 455 m/s. No Splitting
Enabled
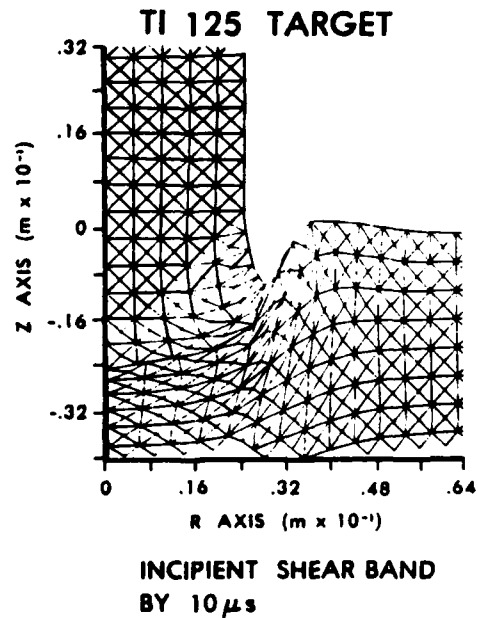
Figure 3. Impact of Steel Cylinder/
Ti 125 Target @ 350 m/s. No Splitting
Enabled

III. APPROACH TO MODELING ADIABATIC SHEAR

A. Constitutive Equations

Prior to enabling splitting between elements, the impact of a steel cylinder against a Ti318 (6%Al, 4%V) target at a striking velocity of 455 m/s was simulated (Figure 2). The Ti318 material is known to fail by adiabatic shear. Contrast these results with those for the impact of a similar steel cylinder against a Ti125 (99% pure Ti) target at 350 m/s (Figure 3). Although the first case involved a much higher striking velocity, there is no sign of an incipient shear band by 10 μs. In the second case plugging failure is due to high strains and an incipient shear band is demonstrated by 10 μs. Such diverse results, and indeed an attempt with the same simulation techniques, indicated that adiabatic shear can not be successfully modeled utilizing only a linear elastic/plastic stress/strain relationship and an equivalent strain criterion, the combination which succeeded in modeling plugging due to high strains. Ti318 has a considerably lower work hardening rate, a much higher yield strength, and a higher thermal softening rate than Ti125 as evidenced by material tests conducted by Wulf (4).

Recht (5) demonstrated that low values of thermal conductivity, density, specific heat, and work hardening rate and high values of thermal softening rate and yield strength were conducive to adiabatic shear. The author's plan was to take into consideration the temperature rise due to the conversion of plastic deformation into heat.

Johnson (6) suggested expressing stress as the following analytic function of strain, strain rate and temperature for the Ti318 target material:

$$\sigma = (A+B\epsilon^n)(1+C \cdot \ln(\dot{\epsilon}/\dot{\epsilon}_0))K_T$$

where A, B, n, and C are material parameters,

$$K_T = 1 - (T - T_R)/(T_M - T_R),$$

and T, T_R, and T_M are present, room, and melting temperature respectively. Johnson had utilized this form for the shear stress and had reasonable agreement with actual test data for several materials (7). The first

expression is an isothermal relationship between stress and strain. Lindholm (8) had demonstrated that a logarithmic relationship between stress and strain rate held over a range of strain rates from 10^{-4} sec^{-1} to 10^{3} sec^{-1}; hence, the second expression. K_T is a linear thermal softening expression and varies from $K_T = 1$ at room temperature to $K_T = 0$ at melting temperature.

The author utilized data for titanium (6% Al, 4%V) published by Maiden and Green (9). Specifically, curve fitting techniques were applied to the stress/strain curve for $\dot{\epsilon} = 1.5$ $sec^{-1} \cong \dot{\epsilon} = 1$ sec^{-1} to determine the best values for A, B, and n assuming $\dot{\epsilon}_o = 1$ sec^{-1}. C was then calculated by applying the values of A, B, and n to the curve for $\dot{\epsilon} = 20$ sec^{-1}. It was assumed that thermal effects were not involved at such low strain rates; therefore $K_T = 1$ for these determinations.

The values of material parameters A, B, n, and C (0., 1896.2, .095, .0226, respectively) and the melting temperature (1660°C) became user input parameters to this highly modified EPIC-2 code. For an individual target element, strength as a function of strain, strain rate and temperature was utilized only after the yield strength (1029 MPa) was reached, thereby retaining a linear elastic relationship.

The adiabatic temperature rise is given by:

$$\Delta T = \frac{1}{\rho \overline{C}_p} \int_0^{\gamma} \tau d\gamma \qquad (10)$$

where ρ = density,
$\overline{C}_p$ = mean specific heat
τ = stress when the plastic strain is γ.

The author assumed constant stress for the increment of time involved and utilized a product of equivalent stress ($\overline{\sigma}$), equivalent strain rate ($\dot{\overline{\epsilon}}$) and the current time increment for the power expression.

Since estimates vary from 85% to almost 100% as to what proportion of the work of plastic deformation is converted into heat (11,12,13), another user input parameter was the percentage of plastic deformation assumed

converted to heat. The author utilized 100% for the calculations attempted. For the projectile material the standard linear elastic/linear plastic relationship was utilized.

B. Development of Criteria Mandating Target Failure

"If the rate of decrease in strength, resulting from the local increase in temperature, equals or exceeds the rate of increase in strength, due to the effects of strain-hardening, the material will continue to deform locally. This unstable process leads to the catastrophic condition known as adiabatic slip." - Recht(5).

Webster defines "adiabatic" as occurring without loss or gain of heat. Rogers (12) notes that although the term "adiabatic" is misleading since heat is lost to the surroundings at a rapid rate, the heat loss is small relative to heat generation. The velocities involved in typical ballistic impacts result in strain rates of $10^5 - 10^6$ sec^{-1} and the period of time (microseconds) means there is insufficient time for the heat generated by such strain rates to be dissipated by conduction; hence, adiabatic slip or shear.

This, then, is the basis for the criterion of primary importance in attempting to model adiabatic shear - when the rate of thermal softening overcomes the rate of strain hardening, i.e. when $\partial\sigma/\partial\epsilon = 0$.

The following criteria were developed and utilized to simulate the adiabatic shear responsible for plugging failure:

Criterion 1: An element initiates or furthers a split
1) when the rate of thermal softening > rate of work hardening ($\partial\sigma/\partial\epsilon = 0$)
2) when, if furthering a split, it is associated with the "next" node,
3) the magnitude of its shear stress > the magnitudes of its deviatoric axial and radial stresses,
4) the direction of the split (Criterion 3) is not to a node with a higher z coordinate.

Criterion 2: The node at which initial splitting occurs
1) must belong to the element meeting Criterion 1,
2) must suffer the highest force of all three nodes belonging to the same element, and
3) must be a master node.

Criterion 3: The direction of the split is determined by the strain of the element meeting Criterion 1. In general, if the magnitude of the axial strain > the magnitude of the radial strain $|\epsilon_z| > |\epsilon_r|$ splitting is to the

nearer radial node. Otherwise, splitting is to the node closer to the direction of radial strain. For more specifics on this criterion see Reference 14.

IV. CASE I

A. Experimental Results

Woodward fired five shots of hardened, roller-bearing steel cylinders into Ti318 targets at normal obliquity (15). Note the geometric and material properties for the first penetrator and target, Table 1. The critical velocity to perforate the target was 441 ± 14 m/s. In every case the projectile fractured. In one case plugging due to adiabatic shear occurred; the plug was pushed through the target but a small part of the projectile stayed in the target while the remainder rebounded (Figure 4a). However the adiabatic shear band was so narrow that it took metallographic analysis (1000X magnification) to ascertain that it definitely was adiabatic shear. In two other cases the projectile broke up and ricocheted and the plug was not pushed out; the shear band did not reach the rear surface of the target but did precede the crack (Figure 4b).

B. Simulation and Comparison

The calculation to simulate this situation utilized a striking velocity of 455 m/s. The pattern of elements which enabled splitting and the elements which totally failed are shown in Figure 5. The deformation history is shown by axisymmetric blowups of the region of activity in Figure 6. Note that the shear band reached the rear target surface at 3.25 μs implying a crack velocity of 1954 m/s. The plug was completely formed and was pushed approximately 1.5 mm past the rear target surface. The forward motion of the projectile stopped between 25 and 30 μs and the projectile is progressively rebounding at 30 μs and 35 μs. Note that several projectile and target elements were automatically totally failed, primarily on the basis of a minimum time increment violation. There were some problems with overlap of elements on the same sliding surface with so much element failure; surface elements can not interfere with elements on another surface but there is no provision for checking interference between elements on the same surface. The speed history of the projectile is shown in Figure 7; the depth-of-penetration history in Figure 8.

The results of the calculation are promising when compared with the experimental evidence with regard to the shape of the deformed rod, the total failure of part of the projectile, the formation of the plug, even the inward-directed angle toward the bottom of the incompletely- formed plug when the velocity is just under the limit velocity. The jaggedness of the splitting pattern would seem to detract from the results of the

MATERIAL	PENETRATOR 1 STEEL	TARGET 1 Ti:318 (Ti WITH 6%Al, 4%V)	PENETRATOR 2 STEEL	TARGET 2 Ti:318 (Ti WITH 6%Al, 4%V)
SHAPE	BLUNT	50mm SQUARE	BLUNT	50mm SQUARE
MASS	3.34g	—	3.34g	—
LENGTH	25.4mm	THICK $6.35mm$	25.4mm	THICK $6mm$
DIAMETER	4.76mm	—	4.76mm	—
DENSITY	7.39×10^3kg/m³	4.43×10^3kg/m³	7.39×10^3kg/m³	4.43×10^3kg/m³
σ_Y	$2290MPa$	1029MPa	$2233MPa$	1029MPa
σ_U	2500MPa	1209MPa	2500MPa	1209MPa
ϵ_U	—	0.2	—	0.2
SPECIFIC HEAT	—	5.65×10^2J/kg°C	—	5.65×10^2J/kg°C
YOUNGS MOD	1.93×10^5MPa	1.158×10^5MPa	1.93×10^5MPa	1.158×10^5MPa

Table 1. Geometric & Material Properties – Case I & Case II
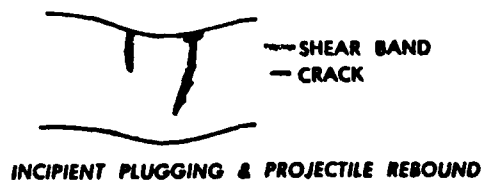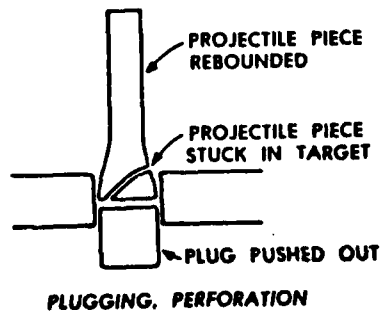
Figure 4a. Sketch of Experimental Figure 4b. Sketch of Incipient Plugging
Results - Case I $V_S < V_L$
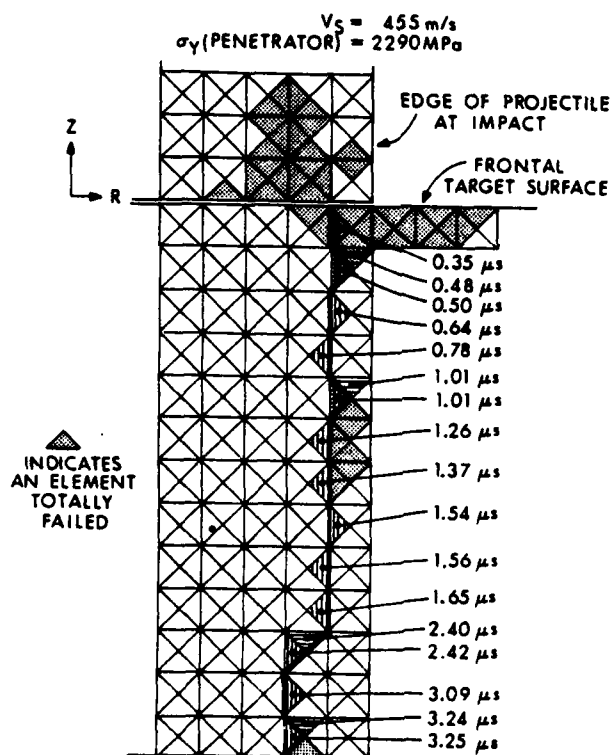
Figure 5. Pattern of Elements Enabling
Split - Case I
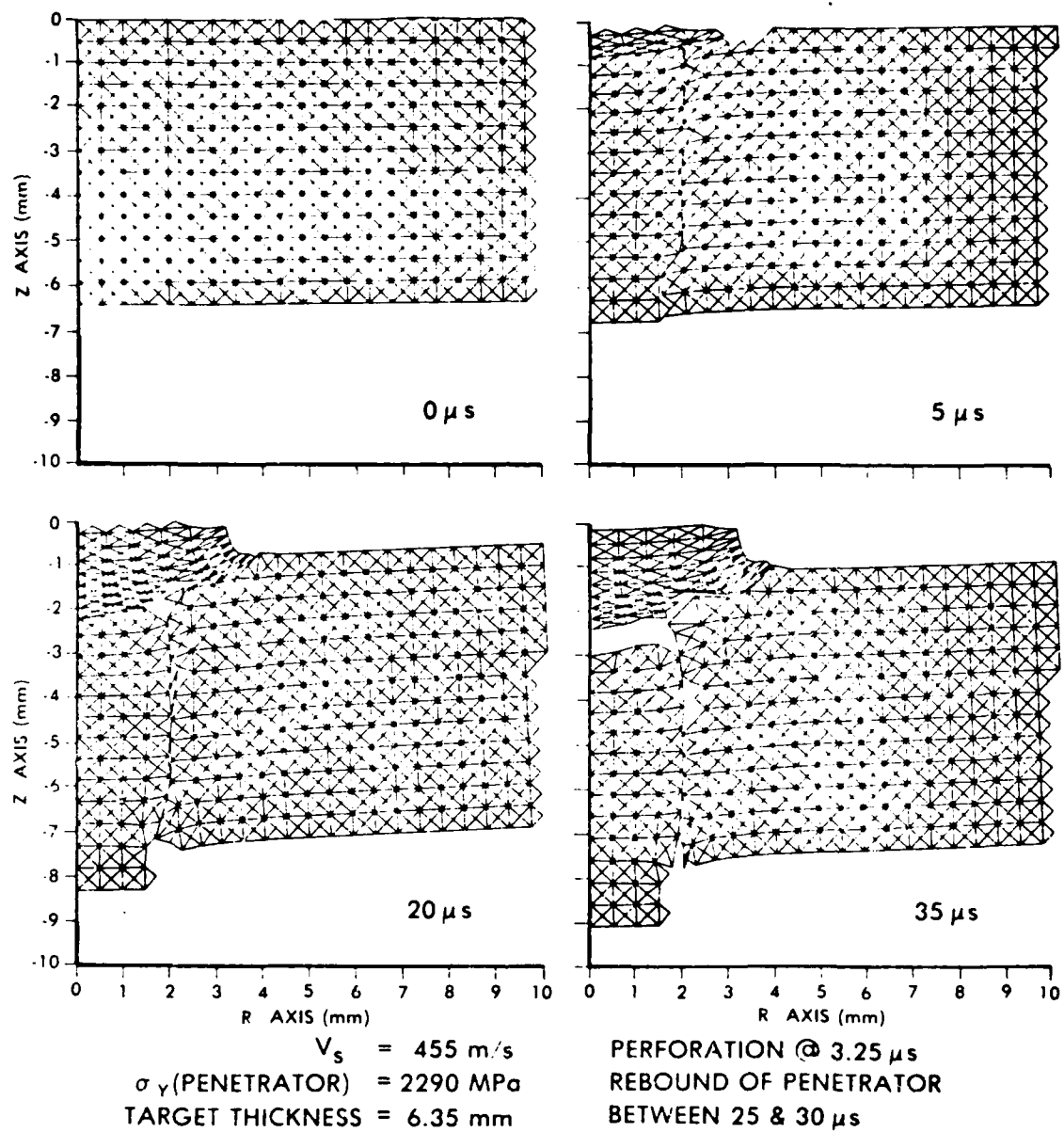
Figure 6. Axisymmetric Blowup of Region of Activity - Case I
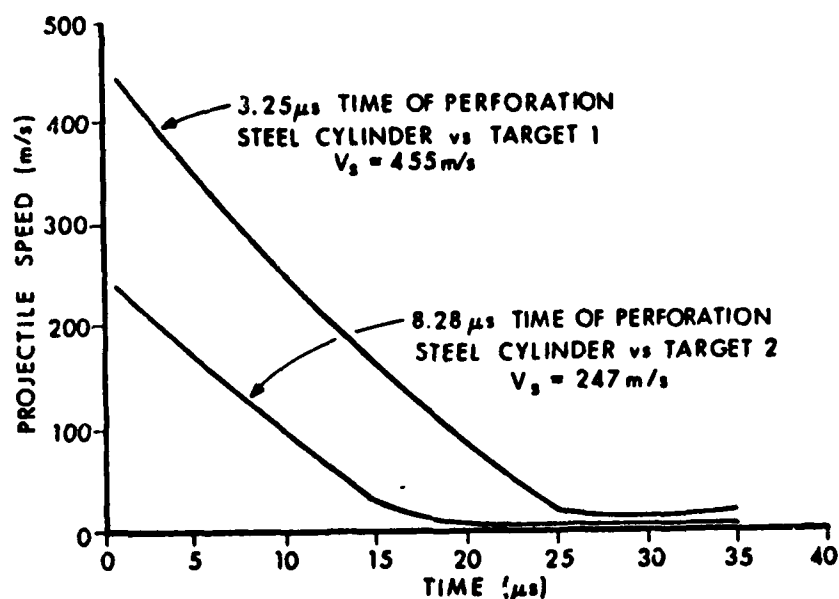
Figure 7. Speed Histories For Case I &
Case II Simulations
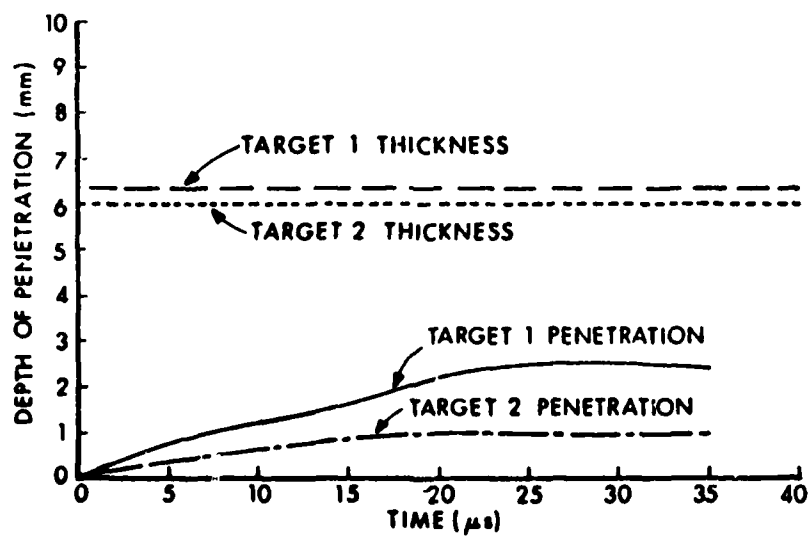
Figure 8. Depth of Penetration Histories
for Case I & Case II

simulation. With splitting disabled, a calculation was made utilizing the stress/(strain, strain rate, temperature) relationship; the instability $\partial\sigma/\partial\epsilon$ = 0 worked its way down through a "bandwidth" of elements; it was not as localized as would be expected. Of more concern is the placement of the split, not directly under the edge of the projectile. This may indicate that compression ought to be allowed to proceed for a longer period of time before the shear band starts to form.

V. CASE II.

A. Experimental Results

The second case for which a simulation was calculated was experimental work only recently conducted by Woodward, Baxter, and Scarlett(16). A blunt steel cylinder with corners ground to a radius of .25 mm impacted a Ti318 target at normal obliquity with a striking velocity of 247 m/s. Results (Figure 9) indicate very little plug compression and the separation of a plug from the remaining target material at several places, but not every-where, along the band. Woodward, et al, noted that examinations of sections parallel to the plane of the plate showed the separate nucleation of very narrow shear bands and their independent propagation to the rear surface despite joining circumferentially in several positions. They also noted an initial acceleration phase during which compression takes place preceding the formation of the adiabatic shear band.

This case differed from the first case in the significantly lower striking velocity (247 vs 455 m/s), the lower hardness of the projectile (2233 MPa vs 2290 MPa), and the different target thickness (6 mm vs 6.35 mm).

B. Simulation and Comparison

This simulation utilized a striking velocity of 247 m/s and the geometric and material properties noted for the second penetrator and target in Table 1. A very jagged splitting pattern is evident in Figure 10. Only two projectile elements were totally failed due to a minimum time increment violation before projectile rebound occurred. Axisymmetric blowups of the region of activity detail the deformation history in Figure 11. The shear band in this case did not reach the rear target surface until 8.28 μs, implying a crack velocity of 725 m/s. A plug was completely formed and extends slightly ($\simeq$ .3 mm) from the rear of the target. There was no discernible plug compression; the projectile began to rebound by 20 μs. The speed and depth-of-penetration histories of this simulation are also shown in Figures 7 and 8.

Figure 9.   Section of Ti 318 - Case II(16)

$V_S = 247 \, m/s$
$\sigma_Y(\text{PENETRATOR}) = 2233 \, MPa$
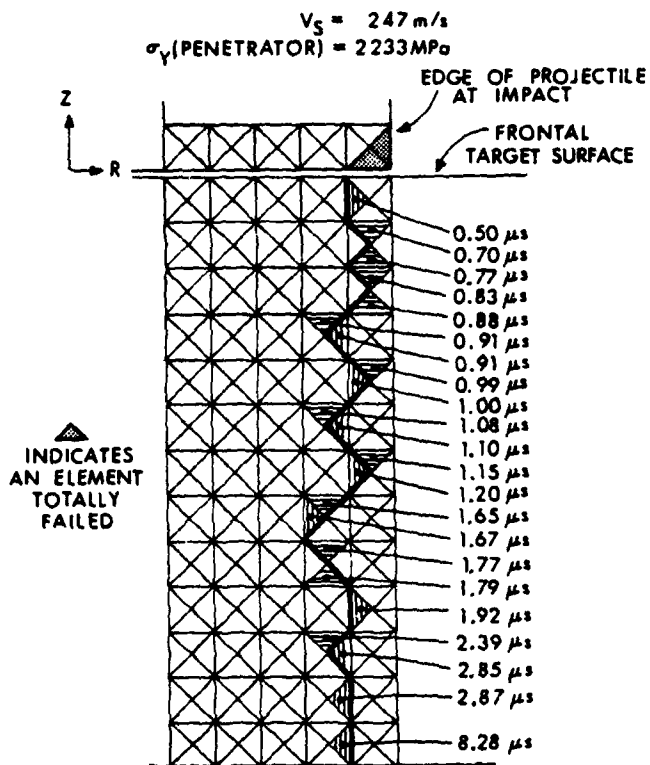
Figure 10.   Pattern of Elements Enabling
Split - Case II
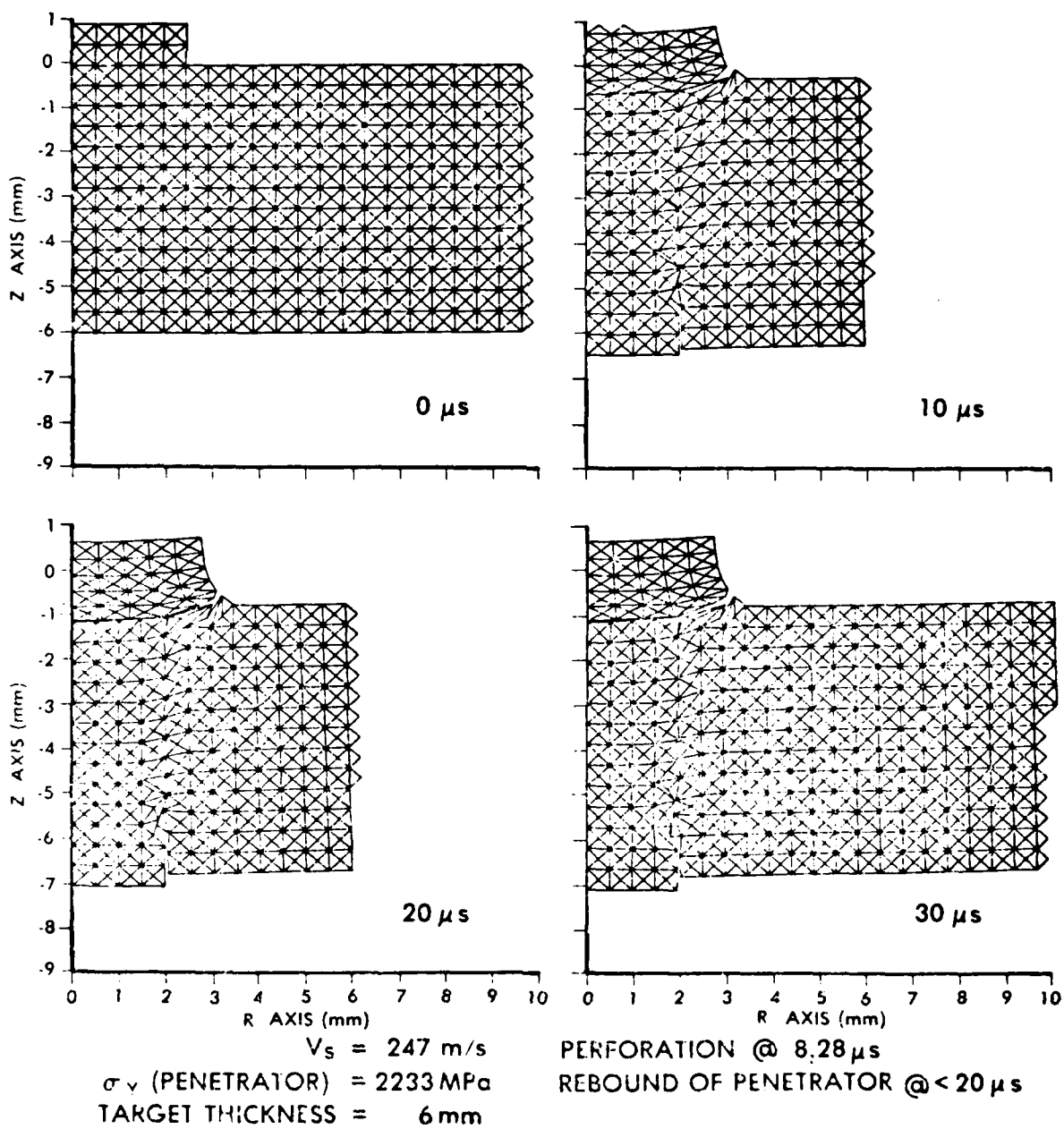
Figure 11. Axisymmetric Blowup of Region of Activity - Case II

RINGERS

Comparison between the more detailed experimental results in this case
and the simulation is most interesting - the correspondingly slower for-
mation of a plug which was not ejected, the highly jagged pattern in the
simulation coincidentally reflecting the areas of separation and
nonseparation of plug from the remaining target. Again the placement of
the split produces a narrower plug than is realistic. The small, but
discernible, compression of the plug and the significant indentation of the
target in the experiment is not matched in the simulation; thus further
supporting the need for a longer compression time.

VI. CONCLUSIONS

This paper presented simulations of target plugging failure due to
adiabatic shear utilizing finite element techniques, a stress/(strain,
strain rate, temperature) relationship, and the monitoring of the
instability which occurs when thermal softening overwhelms work hardening.
Two simulations and comparisons with experiment indicate considerable
promise for the method; results were consistent between calculations - a
much higher striking velocity producing and dislodging a plug whereas the
lower velocity produced a plug more slowly and which remained tightly
lodged in the target. Investigation is planned to determine the necessary
modifications to enable a longer plug acceleration period before shear
failure occurs, again, on natural physical bases.

As indicated by Rogers (12) and Woodward, et al, (16) the adiabatic
shear band forms entirely before there is actual separation. A more
accurate simulation may result when the stress/(strain, strain rate,
temperature) relationship is utilized in conjunction with the original
equivalent strain criterion for actual separation.

It is imperative that criteria designed to model target failure modes
be incorporated with available simulation techniques. Simulations which can
handle only partial penetration of a target can betray penetrator and armor
designers. The results shown in Figures 2 and 3 would seem to imply that
the Ti318 material is tougher than the Ti125 material when, in truth, the
Ti318 material is easier to defeat by kinetic energy penetrators. Ti318
has a much higher static strength but its failure mechanism, adiabatic
shear, requires far less energy (17). The true ballistic worth of these
techniques will soon be ascertained when they are applied to impacts
involving rolled homogeneous armor and high hard armor.

126

REFERENCES

1. Johnson, Gordon R., "EPIC-2, A Computer Program for Elastic-Plastic Impact Computations in 2 Dimensions Plus Spin, "US Army BRL Contract Report ARBRL-CR- 00373, June, 1978.

2. Ringers, B. E., "New Sliding Surface Techniques Enable Lagrangian Code to Handle Deep Target Penetration/Perforation Problems", Computational Aspects of Penetration Mechanics, Lecture Notes In Engineering, Springer-Verlag, 1983.

3. Ringers, B. E., "Simulations of Ballistic Impact Situations Involving Deep Penetration and Perforation of Targets With A Lagrangian Impact Code", Proceedings of the Army Symposium on Solid Mechanics, 1982 - Critical Mechanics Problems in Systems Design, AMMRC MS 82-4, September, 1982.

4. Wulf, G. L., "High Strain Rate Compression of Titanium and Some Titanium Alloys", Int. J. Mech. Sci, Vol. 21.

5. Recht, R. F., "Catastrophic Thermoplastic Shear," Journal of Applied Mechanics, June, 1964, pp. 189-193.

6. Johnson, Gordon R., private communication.

7. Johnson, G. R., Hoegfeldt, J. M., Lindholm, U. S., and Nagy, A., "Response of Various Metals to Large Torsional Strains Over A Large Range of Strain Rates - Part 2; Less Ductile Metals, "Journal of Engineering Materials and Technology, January, 1983, p. 48.

8. Lindholm, U.S., "Some Experiments with the Split Hopkinson Pressure Bar, "J. Mech. Phys. Solids, Vol. 12, 1964, p. 317.

9. Maiden, C. J., Green, S. J., "Compressive Strain-Rate Tests on Six Selected Materials at Strain Rates From 10^{-3} to 10^4 In/In/Sec, "Transactions of the ASME, September, 1966, p. 496.

10. Campbell, J. D. and Ferguson, W. G., "The Temperature and Strain-Rate Dependence of the Shear Strength of Mild Steel," Phil. Mag., Vol. 21, 1970, p. 63.

11. Johnson, W., Impact Strength of Materials, Edward Arnold Ltd., 1972.

12. Rogers, Harry C., "Adiabatic Plastic Deformation," Ann. Rev. Mater. Sci., Vol. 9, 1979, p. 283.

13. Titchener, A. L. and Bever, M. B., Prog. Metal Phys., Vol. 7, 1958.

14. Ringers, B. E. "Modeling Adiabatic Shear", BRL Technical Report, to be published, 1984.

15. Woodward, R. L., private communication - experiments performed at Materials Research Laboratories, Australia.

16. Woodward, R. L., Baxter, B. J. and Scarlett, N. V. Y., "Mechanisms of Adiabatic Shear Plugging Failure in High Strength Aluminum and Titanium Alloys", 3rd International Conference on Mechanical Properties of Materials At High Rates of Strain", Oxford, April, 1984.

17. Woodward, R. L., "The Penetration of Metal Targets Which Fail By Adiabatic ShearPlugging", Int. J. Mech. Sci., Vol. 20.

ANALYSIS OF WORLD AREAS TO INCREASE ARMY MOBILITY (U)

JAMES H. ROBINSON, JERRY L. SMITH, AND GERALD W. TURNAGE
U. S. ARMY ENGINEER WATERWAYS EXPERIMENT STATION
VICKSBURG, MISSISSIPPI 39180

Introduction

For the past several years, the U. S. Army Engineer Waterways Experiment Station (WES) has been developing procedures to assemble 1:50,000-scale mobility-terrain data for use with the Army Mobility Model (AMM) (1-3). The methods developed by WES use multiple unclassified sources of information, primarily maps and collateral information not initially in map form, which are analyzed and consolidated into maps. For a given area, five to ten maps containing information on soils, slopes, land use, etc., are input to a computer in digital form, rectified in the computer to a common scale and coordinate system, and assigned appropriate attribute codes for all areas and/or linear elements, depending on map content. The full map set is then overlaid in the computer to create a composite map as an array of rectangular cells (usually 100 by 100 m for areal features, 10 by 10 m for linear), with attributes from all input sources assigned to each cell.

The total available information for each cell--some quantitative but most descriptive--is interpreted in the computer, using the latest WES artificial intelligence interpretation routines, to assign quantitative values for all mobility-terrain factors needed by AMM (4). The overall accuracy of WES mobility-terrain data has been steadily increased by use of more and better source information; acquisition of appropriate, extensive ground truth data when feasible; use of more airphoto information; and continuous improvement in the artificial intelligence algorithms. Still, the current mobility-terrain data base product is considered to be of "study" quality only. That is, a map developed by this process is a realistic representation of the distributions severities, associations, and general placement of mobility impediments to be found in the area, but no claim can yet be made to high accuracy in the quantitative values for all mapped factors at specific map coordinates. This study level of mapping realism, however, has proved sufficient for successful use of AMM in a wide range of vehicle design, candidate vehicle selection, and

computer war game applications where, in each case, land areas covering one or a small number of 1:50,000 scale topographic map sheets (quads) are considered.

The Problem

Many world areas are of potential interest for AMM applications. One such area is the Federal Republic of Germany for which WES has developed areal mobility-terrain data for 22 full 1:50,000-scale quad sheets and five partial quad sheets (of a country-wide total of some 500 quads) (Figure 1). The cost to map one quad sheet by current research methods is nearly $50,000. For AMM "study" applications, it is considered too expensive to map all of the Federal Republic of Germany (or any area in the world of similar size and complex surface characteristics) in terms of AMM mobility-terrain factor values. Users of AMM, however, require similar mapped data for a number of such large areas worldwide. The challenge, then, is to represent mobility problems over a large area at a reasonable cost.

The Approach

In response to expressed user needs, WES has developed and begun implementation of an approach based on breaking up large areas into a modest number of regions and selecting within each region one or a small number of subareas that are statistically representative samples of the whole from a mobility perspective. The novel aspect of the WES method is that it is systematically applied and computerized and uses objective decision criteria.

Clearly, if quantitative mobility-terrain data were available for all of an area of concern, selecting small representative areas would be merely a massive computer exercise using largely available algorithms. And the study might not even be needed. The problem, rather, is to identify sample areas based on available unclassified, small-scale information not directly addressing ground mobility; the type of information found in national atlases, agricultural studies, and climatic summaries.

The WES method was developed first for the total 250,000 sq km of the Federal Republic of Germany. This was because of the military interest in this area, the reliability of plentiful available atlas-type data for direct use and for follow-on verification, the accessibility of the area for ground truthing, WES's familiarity with the area on the ground, and the support of the WES effort by the Republic's Ministry of Defense.

Regionalization for the Study

Ground mobility bears a close relation to the detailed configuration and materials of an area. To begin regionalization, it was assumed that
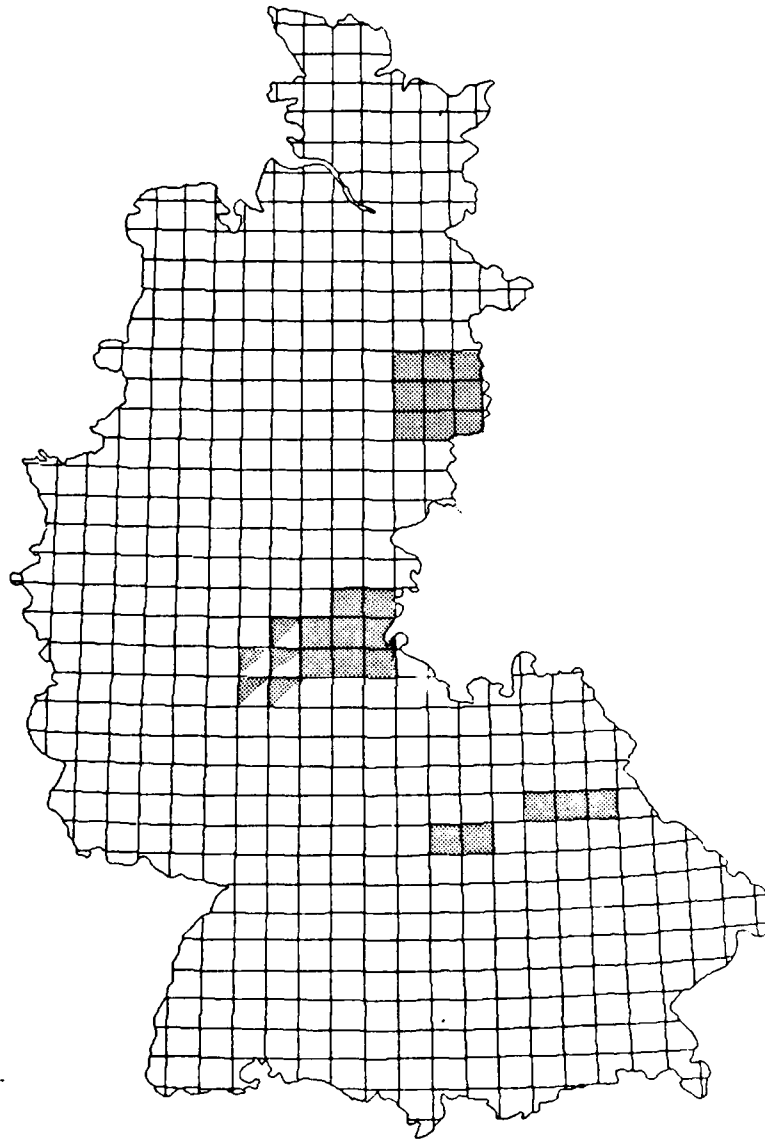
Figure 1. WES coverage of the Federal Republic of Germany
in terms of AMM mobility-terrain factors (by 1:50,000-
scale quad sheets)

131

landform and climate were the overriding influences on the regional terrain details that ultimately determine vehicle performance.

Landform was designated by one set of large patches described in five broad classical geographic classes, and by a second mosaic of patches described by 16 classes of characteristic slope numbers derived from a 1974 Natick study (5) (Figure 2).

Two simple long-term climatic factors were used: mean annual rainfall and average January air temperature (AJAT) (Figure 2). Rainfall affects erosion, soil moisture, vegetation growth, etc., and thus indirectly the terrain factor values influencing vehicle performance. Rainfall boundaries were selected following Holdridge's Life Zone concept (6). AJAT in the Federal Republic of Germany was shown by Billelo and Appel (7) to be generally related to snowfall, length of growing season, and temperature patterns throughout the whole year--all of consequence to the more detailed mobility-terrain factors of interest.

When maps of these four factors were overlaid, the total land area became a mosaic of some 200 regional patches, each described by the four factors. Patches covering a country-wide area of less than 1 percent of the total (approximately five 1:50,000 quad sheets) were considered too small to constitute a region. When these were eliminated, there remained 23 regions covering 70 percent of the entire country (30 percent of the country fell into the undifferentiated category). Figure 3 shows the way in which the 23 regions and undifferentiated areas break up the country. (The map symbology does not imply that all areas with the same symbol are in one region.) Figure 3 is at a resolution of 10- by 10-km cells or sample cells; each cell covers approximately the area of one fourth of a standard 1:50,000 quad.

Finding Sample Areas

Unclassified atlas-type, small-scale maps depicting five more-detailed aspects of the countryside were assembled and input for analysis in the computer as 400- by 400-m data cells. Data for the five aspects are as follows:

Terrain Aspect	Final Weighting	No. of Classes
1. Unified Soil Classification System (USCS) soil type	4	10
2. Agricultural land use	3	17
3. Forestry type	2	3
4. Vegetation cover (types)	1	9
5. Geology (rock types)	1	8

Topography

Factor 1

Slope Categorization (from Natick Landform Classification System)

Classes*

(1)	0-2%	(9)	0-2%
(2)	2.1-5%	(10)	2.1-5%
(3)	5.1-10%	(11)	5.1-10%
(4)	10.1-20%	(12)	10.1-20%
(5)	20.1-40%	(13)	20.1-40%
(6)	40.1-60%	(14)	40.1-60%
(7)	60.1-70%	(15)	60.1-70%
(8)	>70%		>70%

Factor 2

Landform-Topography (Landform, Slope, Relief)

Classes

Landform Types	Average Slope, %	Relief, m
(1) Marsh	<3	<5
(2) Level plains	3 to 8	5 to 25
(3) Rolling plains and hills	8 to 20	25 to 50
(4) Hills	20 to 30	50 to 100
(5) Mountains	>30	>300

Climate

Factor 3

Avg Jan Air Temp, $^{\circ}C$

Classes

(1) >0
(2) 0 to -1
(3) <-1 to -2
(4) <-2

Factor 4

Annual Rainfall, in.

Classes

(1) >20 to 40
(2) >40 to 80

* Classes 1-8 and 9-16 are based on single and dual classification systems, respectively.

Figure 2. Four primary factors presently used in WES regionalization of the Federal Republic of Germany (from thematic maps at scales around 1:1,000,000) (To convert inches to centimetres, multiply by 2.54.)

133
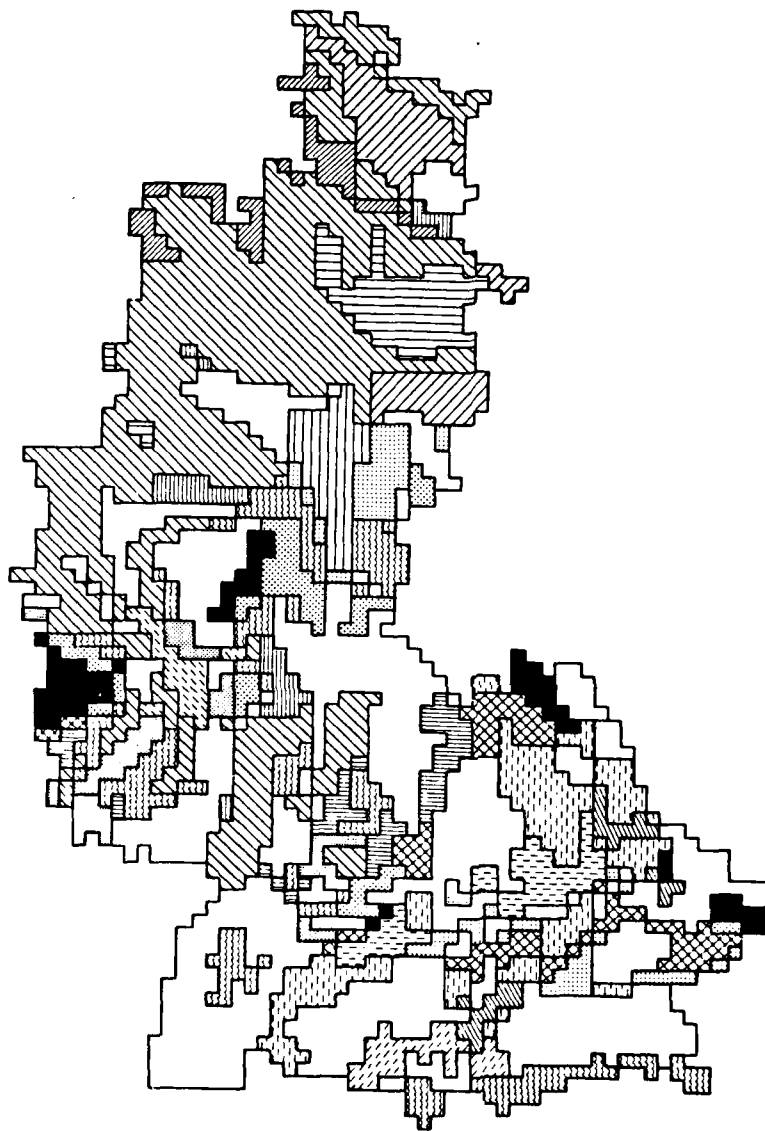
Figure 3.   Twenty-three regions covering 70 percent of the
Federal Republic of Germany

Again, none of these were the quantitative mobility-terrain factors needed for use in AMM, but they were representative of the higher resolution information used by WES to develop detailed AMM data bases at 1:50,000 scale.

The 10- by 10-km grid net used in regionalizing was superimposed on the five aspect maps and simple distributions of each aspect determined for each sample cell, and for all sample cells throughout each region. For each of the five aspects, correspondence of the distribution of class descriptors within each sample cell to the region-wide distribution was made on the basis of the chi-squared (χ^2) statistic (8-11). For a single terrain aspect described in m classes

$$\chi^2 = \sum_{i=1}^{m} (o_i^2 - e_i^2)/e_i$$

where

o_i = the observed value for class i

e_i = the expected value for class i

The χ^2 process tests to some level of confidence the hypothesis that a given distribution (such as for a sample cell) is identical to another (such as for the whole region). The formal test becomes very stringent when the sample cell distribution derives from many "tests." In the present application, using data from 400- by 400-m data cells over a 10- by 10-km sample cell results in 625 tests. Results of formal testing did not prove useful. What was needed was a means to determine some degree of closeness other than full matching. To this end, the value of χ^2 for a specific sample cell was tested against the standard deviation of χ^2 for all cells in the region to give a Z score:

$$Z \text{ score} = \frac{\chi^2 - \bar{\chi}^2}{STD_{\chi^2}}$$

where

χ^2 = the value for a specific cell in a region

$\bar{\chi}^2$ = the mean value of χ^2 for all cells within the region

STD_{χ^2} = standard deviation of the distribution of χ^2 for all cells within the region

Z_{score} provides a normalized description (in terms of standard deviations) of the variation of the χ^2 of a specified cell relative to the χ^2 distribution of all other cells within a region. The smaller the Z_{score} value, the more nearly the distribution of the cell represents the distribution of the region.

Z_{score} values were computed for all five aspects in all sample cells, coded for degree of acceptability (dot = <-0.5, closed diagonal = >-0.35 to 0.25, and open diagonal >0.25) and printed in map form for inspection. Figure 4 is an example for one aspect (USCS soil type) throughout one large region. In general, most unacceptable sample cells--in all such presentations--tended to be near region boundaries. This was not unexpected. Unacceptable sample cells embedded within a region proved, on inspection of the relevant topographic map sheet, to have clearly anomalous features, which also was not distressing.

Unacceptable cells were eliminated wherever found--tightening the regional characterization--and the process repeated up to three times, at which point the pictures pretty well stabilized. The final Z_{scores} for the five aspects were combined based upon a judgmental rating of the relative importance to mobility-terrain factors of matching the local distributions of the separate aspect to the regional distributions. While judgmental, these ratings derive from in-depth experience in mobility performance and mapping research. Weightings were as listed earlier. For each cell, weightings were multiplied by the Z_{scores} for the five aspects and the results totaled. The aggregate scores were classed and mapped by sample cells for final decision at this time. Sample cells with the lowest aggregate score were considered most representative of the region.

Figure 5 shows the twenty-three 20- by 20-km areas (square blocks of four sample cells--approximately the size of one 1:50,000 quad sheet) selected by examination of aggregate scores map to represent the 23 current regions. The background grid shows the 1:50,000 topographic sheets for the area.

These 23 areas are prime candidates for the generic mobility-terrain data base. When detailed with mobility-terrain factor values, mobility performance statistics predicted for these areas can be used to study mobility potentials, problems, and options for approximately 65 percent of the land area of the Federal Republic of Germany. (At this moment, it is estimated that temporarily relegating regional boundary transition areas to the undifferentiated category will drop the original coverage by about 5 percent.)

The Future

While Figure 5 represents the first important result of the generic
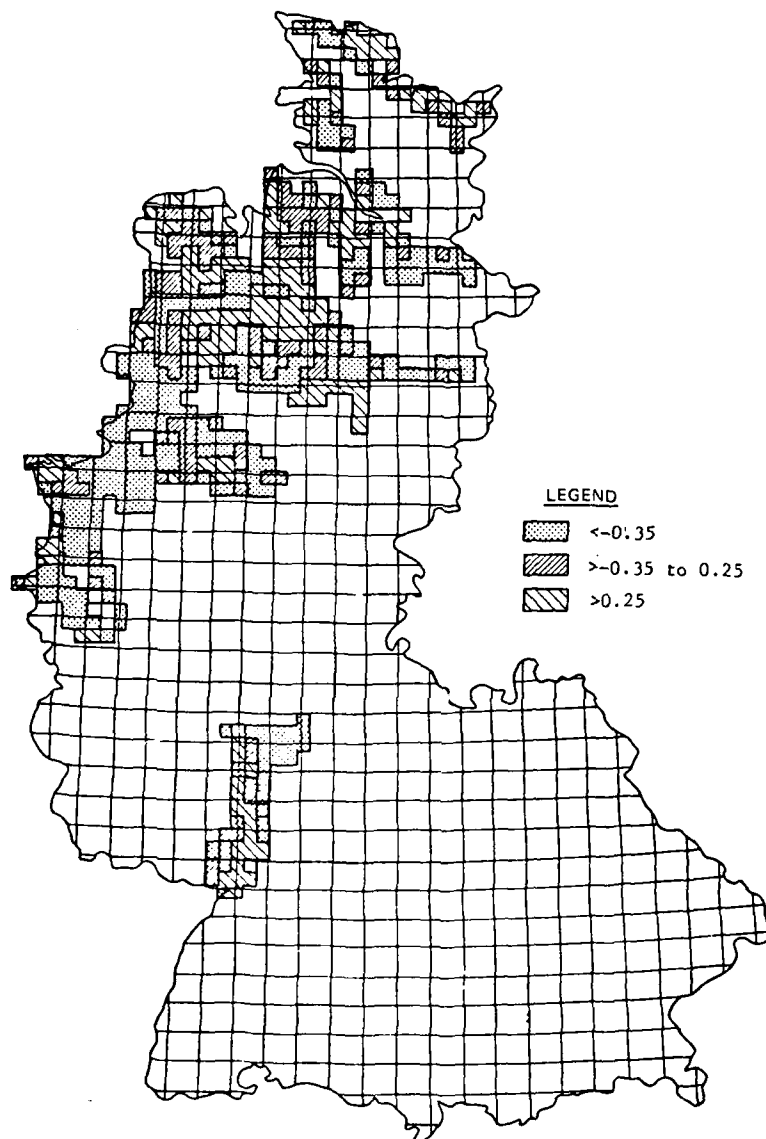
Figure 4. Distribution of Z_{score} values for the chi square statistic (USCS soil type, Region 1, Federal Republic of Germany, 1:4,000,000 scale, 10-km resolution)
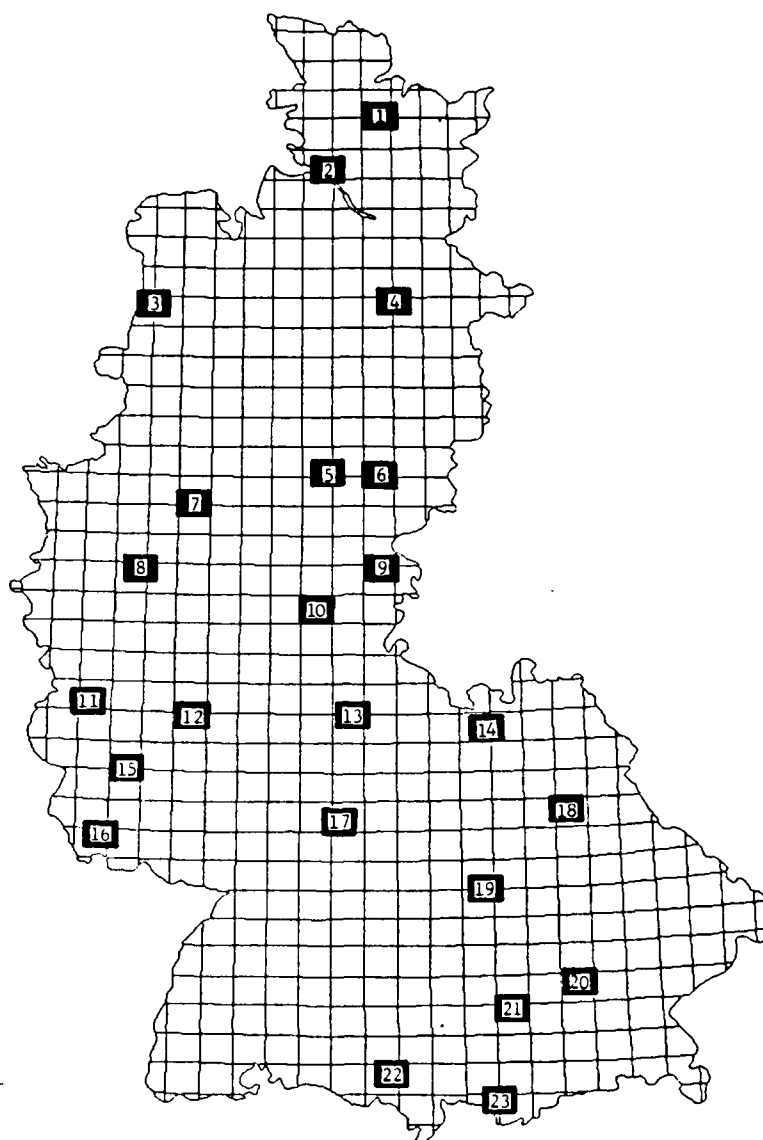
Figure 5. Twenty-three 20- by 20-km areas representing
the 23 regions covering 70 percent of the
Federal Republic of Germany

terrain study, work is still in progress. In the near future, the original regional boundaries must be refined, slope distributions added to the χ^2 analyses, and current ideas and methods extended to classify the presently undifferentiated areas. The regionalization method must be verified. Brief on-the-ground spot checks of early results in summer 1982 revealed no fatal flaws. Strenuous, more formal testing will be done in early fall 1984, in which more than forty 10- by 10-km sample cells will be characterized by ground measurements. Cells will be selected insofar as possible to check both homogeneity within regions and differences among regions, and to collect data to test emerging ideas about characterizing the presently undifferentiated areas.

The final immediate concern is that, before the generic data base can become really useful, detailing of the 23 areas to quantitative mobility-terrain factor values must be accomplished. WES efforts in this direction will begin concurrently with the field verification work, by using verification cell data as part of the needed quad sheet area ground truth data.

Farther down the road are trial applications over other large areas, and similar treatment of road net and gap mobility-terrain factors.

A Closing Note

WES has deliberately taken a formalized approach to regionalization with the goal of developing a method that can be applied to other large areas of interest anywhere in the world. The Federal Republic of Germany might well have been regionalized by many other detailed approaches because of the wealth of reliable data available on almost any subject except mobility per se. But the Federal Republic of Germany is a special case. Accordingly, WES has attempted to develop a process based primarily on relatively simple, low-resolution information likely to be available for almost any large area, and to use the more detailed information on the Federal Republic of Germany to test hypotheses and results during this research phase. At this stage, it remains to be seen how well worldwide goals have been met, but there is reason for optimism.

References

1. U. S. Army Engineer Waterways Experiment Station and U. S. Army Tank-Automotive Command, Technical Report No. 11789 (LL-143), "The AMC '71 Mobility Model," June 1973, U. S. Army Tank-Automotive Command, Warren, Mich.

2. Jurkat, M. P., Nuttall, C. J. Jr., and Haley, P., "AMC-74 Mobility Model," Technical Report No. 11921 (LL-149), May 1975, U. S. Army Tank-Automotive Command, Warren, Mich.

3. Nuttall, C. J. Jr., and Randolph, D. D., "Mobility Analyses of Standard- and High-Mobility Tactical Support Vehicles (HIMO Study)," Technical Report M-76-3, February 1976, U. S. Army Engineer Waterways Experiment Station, CE, Vicksburg, Miss.

4. West, H. W., Krivitzky, D., and Randolph, D. D., "Digital Terrain and Mobility Data Bases for E-FOSS," Miscellaneous Paper EL-80-1, January 1980, U. S. Army Engineer Waterways Experiment Station, CE, Vicksburg, Miss.

5. Anstey, R. L., "The Natick Landform Classification System," Technical Report No. 100, October 1974, U. S. Army Materiel Systems Analysis Activity, Aberdeen Proving Ground, Md.

6. Holdridge, L. R. et al., Forest Environment in Tropical Life Zones: A Pilot Study, Pergamon Press, New York, 1971.

7. Bilello, M. A., and Appel, G. C., "Analysis of the Midwinter Temperature Regime and Snow Occurrence in Germany," Report 78-21, September 1978, U. S. Army Cold Regions Research and Engineering Laboratory, Hanover, N. H.

8. Wine, R. L. 1964. Statistics for Scientists and Engineers, Prentice-Hall, Englewood Cliffs, N. J.

9. Hald, A. 1952. Statistics Theory with Engineering Applications, John Wiley & Sons, New York.

10. Abramowitz, M. and Stagen, I. A. 1965. Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, Dover Publications, New York.

11. Dixon, W. J. and Massey, F. J. Jr. 1957. Introduction to Statistical Analysis, McGraw-Hill, New York.

PHOTOACOUSTIC MICROSCOPY
DOUGLAS N. ROSE, DR.
DARRYL C. BRYK, MR.
US ARMY TANK-AUTOMOTIVE COMMAND
WARREN, MICHIGAN 48090

INTRODUCTION

The photoacoustic effect was first reported by Alexander Graham
Bell in 1880. Within a year, further investigations were carried out by
the famous, including Bell and his associate Tainter, Roentgen and Lord
Rayleigh as well as by the not so famous including Mercadier and also
Peerce who came closest to our modern explanation of the effect. It
then lay completely dormant for 50 years until the advent of improved
microphones allowed quantification of sound. Then it was revived by
Viengerov in the Soviet Union for work with gases. It was not reapplied
to solids until 10 years ago (1). The first photoacoustic microscopy
image was produced less than 5 years ago using a borrowed laser and an
O-ring for a drive belt between a motor and a stage micrometer (2).

PHOTOACOUSTIC PRINCIPLE

One of Bell's early experiments illustrates the effect in solids
well. He had a closed glass vessel (reputed to be a bell jar) with
solid material inside to which a listening tube was attached. When the
glass vessel was illuminated with rapidly interrupted sunlight, a sound
was heard at the interruption or "chopping" frequency. When the
sunshine was stronger, the sound was stronger. Also, the sound was
stronger when the solid material inside had a darker color.

Figure 1 describes the effect schematically. Since the temperature
that the periodically heated spot on the surface reaches depends not
only on optical absorption but also on thermal properties, which
includes thermal boundary conditions, we can probe beneath opaque
surfaces. The highly damped thermal wave that is generated into the
solid gives this field its other name, thermal wave imaging. Figure 2
shows a schematic layout of a modern system utilizing gas cell
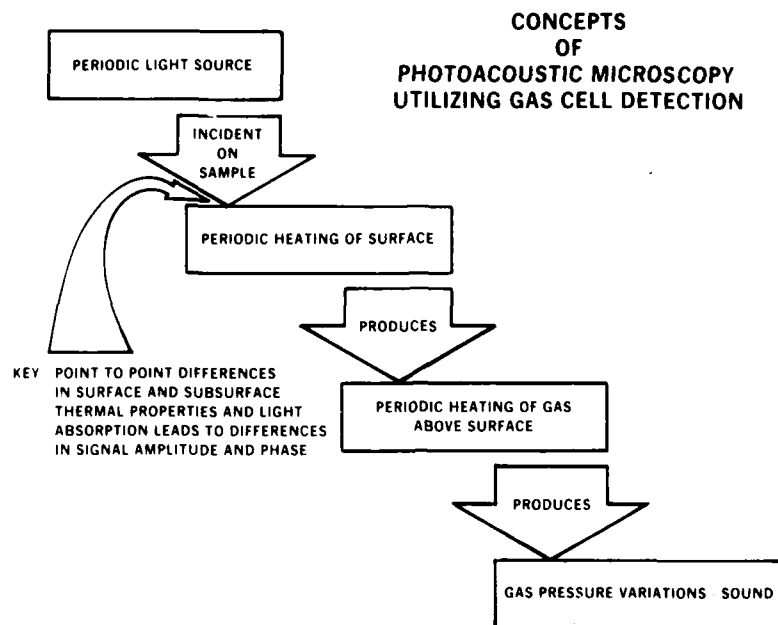
Figure 1. Principle of the photoacoustic effect using gas cell detection.
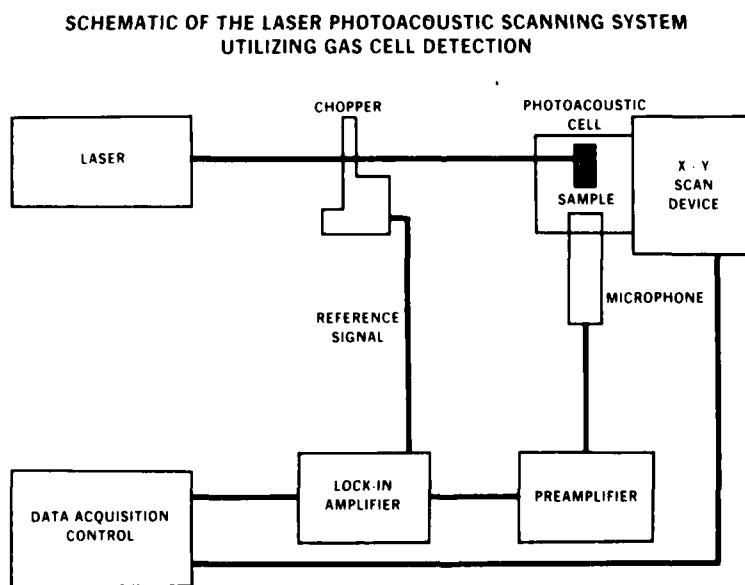
Figure 2. Schematic of photoacoustic microscope using gas cell detection.

detection. The laser is simply a convenient source of directed light; indeed, electron beams (3) and ion beams (4) can be used for spot excitation. Some means of modulating the beam is necessary. Sophisticated electronics such as the lock-in amplifier along with a microphone or other detectors allow us to extract and quantify the signal in a way that was unavailable in the nineteenth century. Figure 3 shows a modern gas cell. Note the window for the radiation to pass through to the surface and the hearing aid microphone. The window is placed over the surface to be examined. There is a shallow channel leading to the hearing aid microphone. The cell is sealed to the surface to be examined with a bead of sticky wax.

The photoacoustic signal has a phase with respect to the excitation beam as well as a magnitude. Good examples of the use of the phase for depth profiling come from the field of photoacoustic spectroscopy in which the interest lies in the variation of the photoacoustic signal with respect to excitation light frequency as the name suggests. For instance, in the spectra of an apple peal (5), the red color shows up in spectra taken with a phase delay. An apple has a natural coating of wax and it takes some time for heat generated in the underlying pigmented layer to become evident at the surface. Similarly, a lobster shell which appears black optically was shown to have different colored dyes at different depths because of the different phase angles associated with different absorbed light frequencies (6). The phase is also somewhat dependent on surface profile though it is relatively insensitive to surface reflectivity (7).

Figure 4 shows how another sample was constructed which illustrates the photoacoustic principle further. Two slots were milled in aluminum and came together forming a sharp point. The slots were filled with plastic and the surface was polished smooth. A coat of silver paint was then applied, hiding the structure from optical view. Figure 5 shows the photoacoustic picture at two modulation frequencies. With quite long thermal wavelengths, the point is sharp and geometrical shadowing holds; however, no depth information can be gained from the phase since the point is at "zero" phase. The lower image at 11 Hz is in this regime but the frequency is approaching the regime where wave effects dominate so some blunting of the point is noticeable. In the upper image at 500 Hz, the thermal wave length is now short compared to the feature and the point is again sharp because the thermal wavelength is small compared to the feature.

DETECTION TECHNIQUES

These last images were taken using the mirage detection technique.
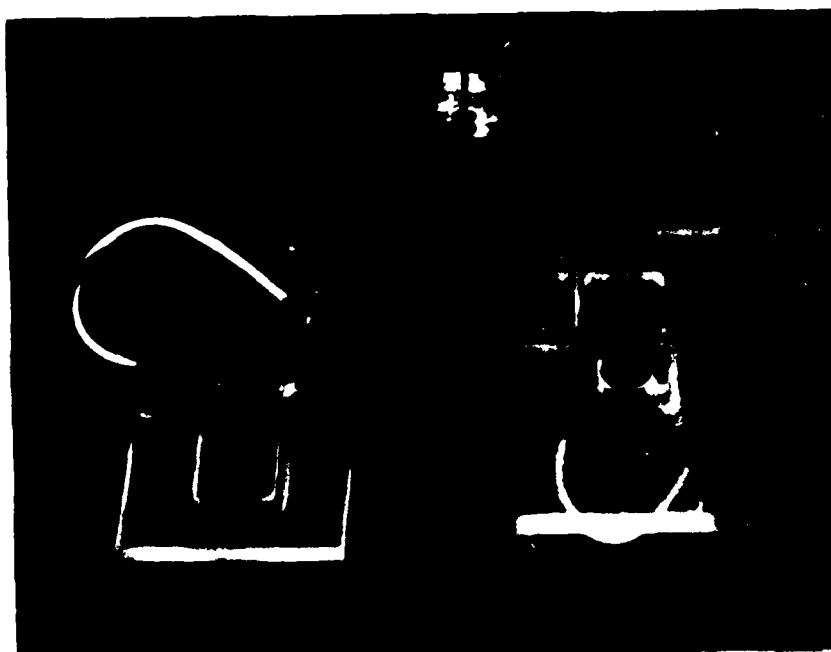
Figure 3. Photoacoustic microscopy gas cell.

CONSTRUCTION OF POINT RESOLUTION SPECIMEN

SLOTS MILLED
IN ALUMINUM

FORMING A
POINT WHERE
THEY MEET

SLOTS FILLED
WITH PLASTIC

AND POLISHED

SILVER PAINT
APPLIED

HIDING POINT
FROM VIEW
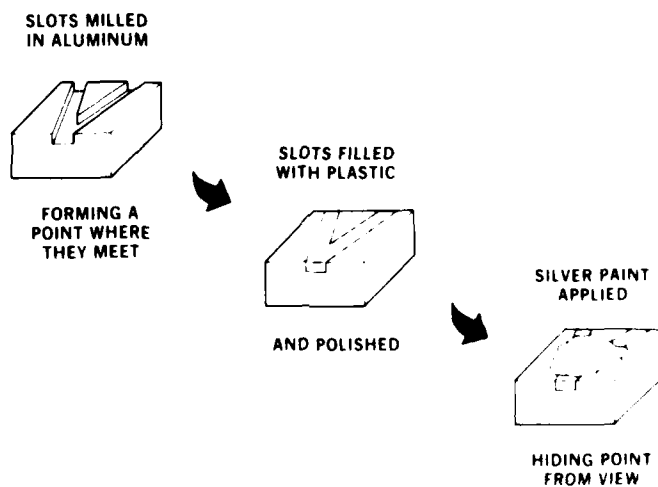
Figure 4. Construction of a photoacoustic point resolution test specimen.
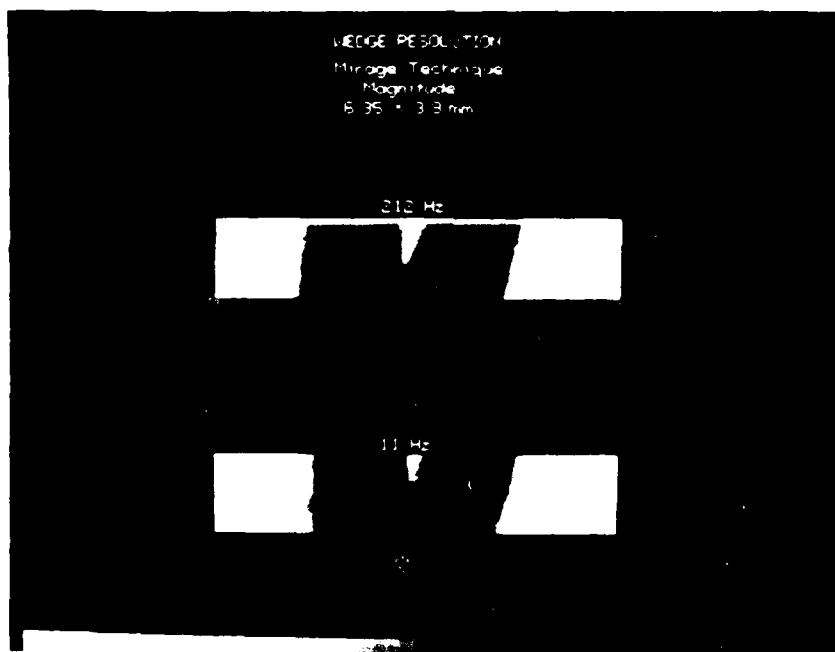
Figure 5. Photoacoustic image of point resolution test specimen.

Figure 6 shows the various effects occurring at the periodically heated spot including gas pressure variations, IR emission and thermal expansion and contraction of the spot. Figure 7 shows 5 of the most common detection techniques used in photoacoustic microscopy. The first one, gas cell detection, is what we now call the approach of Bell which has been described.

The "mirage" technique, also called optical beam deflection (8), was pioneered by the French (9). In this approach, a second laser beam called the probe beam skims within 50 μm of the surface. The periodically heated air above the excited spot deflects the probe beam. The principle is the same as that responsible for what appears to be water ahead of us on a hot highway. With this technique one can go to lower modulation frequencies and thus see deeper with longer thermal

waves. Unfortunately, this makes image acquisition slower because of the longer diffusion time. If the probe beam is slightly displaced from the excitation beam, very characteristic signatures are seen of cracks oriented along the probe beam direction as will be shown. The mirage technique can obviously not be applied to concave surfaces. It also requires maintaining the relative position of the excitation beam focal spot, the probe beam and the sample surface during a scan. With this technique local variations in surface height can produce signal variations in the magnitude and phase that are not easily separated from subsurface thermal wave interactions (10).

Piezoelectric detection relies on the acoustic waves that are generated as the excited spot on the sample periodically expands and contracts. The conversion efficiency from thermal to acoustic energy is typically quite low but this technique can be faster since one does not have to wait for the thermal information to come back to the surface. Thermal wavelengths are typically 0.3 μm in insulators and 3 μm in metals at 1 MHZ dropping to 0.03 μm and 0.3 μm respectively at 100 MHZ. This approach can be used where the exciting beam is a chopped electron or ion beam in which case the sample must be in a vacuum. Using this technique and a modified electron microscope, images have been obtained in the same length of time that it takes to get an electron microscope picture. Subsurface flaws in integrated circuits, the doping pattern in semiconductors, the ion implantation pattern in GaAs and the crystalline structure in welds and also in wire bonds to integrated circuit chips have been imaged in this way (11). This technique does require bonding a transducer to the sample with the usual care necessary for this. Theoretically, the interpretation is more complicated since the material's elastic properties, which are expressed as a tensor, became important and acoustic wave reflections must be considered.

In photothermal imaging, the periodic infrared emissions from the heated spot or the emissions from the back of thin specimens from thermal waves propagating through the specimen are detected. This is a point sampling approach but it relys on thermal diffusion so it is not as inherently fast as piezoelectric detection. It is however, a non-contact method and can also be used in a vacuum. Using this approach, 0.5 μm changes in a 50 μm layer of paint on a 0.5 mm thick sheet of metal have been detected (10) (12). Also the thickness of diffusion hardened surface layers in steel has been measured to an accuracy of 0.15 mm (13). It is adaptable to complex geometries and its resolution is at least 10^{-3} Celsius degrees.

PHOTOACOUSTIC EFFECTS

PERIODIC EXCITATION BEAM

PERIODIC IR EMISSION

PERIODICALLY HEATED GAS

BULGING FROM THERMAL EXPANSION

HEATED VOLUME

ACOUSTIC WAVES IN SOLID
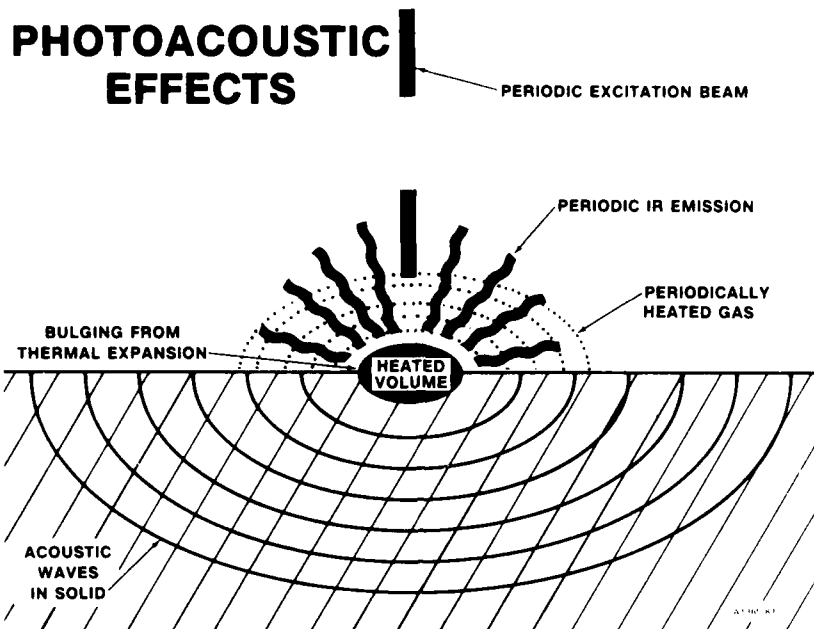
Figure 6. Effects at a periodically heated spot.

PHOTOACOUSTIC MICROSCOPY DETECTION TECHNIQUES

GAS CELL

MIRAGE

PERIODICALLY HEATED GAS

CELL   WINDOW   MICROPHONE

SEAL

SPECIMEN

PHOTODETECTOR

PERIODICALLY HEATED GAS

PROBE LASER BEAM

SPECIMEN

PIEZOELECTRIC

ALTERNATE POSITION OF PIEZOELECTRIC DETECTOR

ACOUSTIC WAVES

SPECIMEN

PIEZOELECTRIC DETECTOR

REFLECTION   PHOTODETECTOR

PERIODICALLY HEATED GAS

PROBE LASER BEAM

PHOTOTHERMAL

PERIODIC IR EMISSSION

IR DETECTOR

SPECIMEN

SPECIMEN

Figure 7. Common photoacoustic detection techniques.

In the reflection approach, a probe laser is reflected directly off the excited spot. It is felt now that changes down to at least 10 nm can be detected (for reference, green light has a wavelength of 550 nm). While this is very good, state of the art ultrasonic transducers can do better. Both thermal lensing, as in the mirage technique, and thermal expansion, as in the piezoelectric approach, affect the signal. This technique does not require a particularly well polished surface. It does require maintaining the relative positions of the excitation beam focal spot and the spot at which the probe beam is reflected which is usually slightly displaced from the excitation beam's focal spot. It is a non-contact method and high modulation frequencies are possible giving fine resolution.

APPLICATIONS

In a general sense, we see photoacoustic microscopy as complimentary to other NDE techniques. For example, it is usable quite near the surface where ultrasonics is more difficult to apply and it can be tolerant of complex surfaces on which ultrasonics is difficult to interpret. Compared to radiography which is not as sensitive to light elements such as silicon, nitrogen and carbon found in many ceramics and composites, it is well suited to ceramic surfaces which can withstand powerful excitation beams thus giving better signal to noise ratios and for which surface flaws are very important. On the other hand, photoacoustic microscopy is quite limited in depth capability and is usually slow compared to the other techniques.

Because of its dependence on microscopic thermal properties, as well as light absorption, photoacoustic microscopy offers a view of surfaces that differs from optical or electron microscopy so that considerable inhomogeneity has been revealed by this technique in samples thought to be uniform. Taking advantage of this, photoacoustic microscopy can be used for characterization of surfaces as well as for flaw detection. A potential application of this is for evaluation of rubber uniformity in track pads (/).

Much of our work has been targeted at ceramics. The newer forms of silicon nitride, silicon carbide and zirconia have higher strengths and offer high temperature capabilities for use in engines. The hardness of silicon carbide and silicon nitride make them attractive in wear situations and their lower density and corrosion resistance make them attractive for turbine blades and turbocharger rotors. However, they are still brittle and hence unforgiving and they typically fail at flaws an order of magnitude smaller than for metals.

EXAMPLES

Figure 8 shows a comparison of photoacoustic images of Knoop indentations in silicon carbide taken using the gas cell and the mirage detection techniques. A Knoop indentation is a hardness test made with an elongated diamond point and is like Vickers hardness testing except that a pyramidal diamond point is used in Vickers measurements. A vertical "half-penny" crack is formed beneath the surface as well as a system of lateral cracks. The visible indentation, 140 X 20 μm is somewhat less than half the height and 6% of the width in the field of view in these images. In relative size it corresponds to the faint gray area in the + 45 degree view of the gas cell images. There is a background signal, minimal for the 90 degree images and highest for the 0 degree images, that has been subtracted from the data. The two sets of data have also been processed so that the values shown stretch from the highest to the lowest grey scale in each set. Though the two sets of images are of different indentations, the difference in information returned by the two techniques is shown. According to theory, the effect of the cracks that are shallow compared to the thermal wavelength should show up at - 45 degrees phase angle (14). Accordingly, the smallest indications can be seen in the + 45 degree images and the biggest indications in the - 45 degree images. The residual indication at + 45 degrees for the gas cell is probably due to the surface profile of the indentation in the surface. The gas cell approach, which averages the temperature over the surface, cannot see a perfectly vertical, perfectly closed crack (15). In contrast, the mirage technique does respond to vertical cracks. Corroborating this, the mirage response to its indentation is much bigger at + 45 degrees than the gas cell response at + 45 degrees. The overall structure of the mirage signal data, which is most clearly shown in the - 45 degree image, is a large mound surrounded by a raised ring. This structure is best shown in a graphical perspective plot like Figure 9 and illustrates that a combination of data presentations will be required to effectively communicate all of the features of the data.
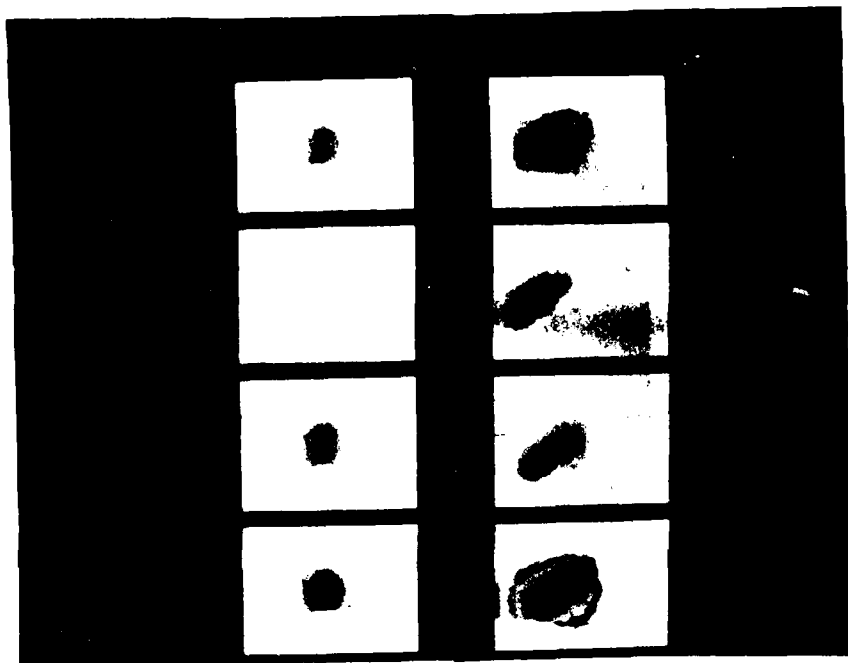
Figure 8. Knoop indentations in sintered silicon carbide: Gas Cell vs Mirage detection.
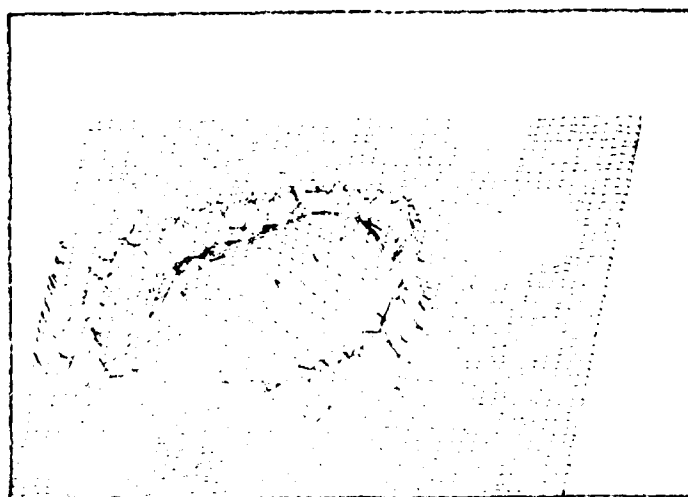
Figure 9. Knoop indentation in sintered silicon carbide: Mirage image at - 45 degrees taken at 212 Hz. The area shown is 318 X 318 microns.

Figure 10 shows a comparison of gas cell and mirage images of a fatigue crack in aluminum that has been stretched open. As can be seen, the mirage technique gives a very distinctive signature for cracks parallel to the probe beam when the beam is slightly displaced from the excitation beam. When the probe beam is on the opposite side of the crack from the excitation beam, there are additional possibilities (or complications depending on ones view) because there is a transverse to surface as well as a normal to the surface deflection of the probe beam. In the mirage image there is a faint crown-like structure above the crack and a vertical line in the upper half of the image. We first noticed this when using a pseudocolor presentation. We have found our pseudocolor capability very useful in revealing subtle detail in the images which we subsequently verify in the gray scale images.
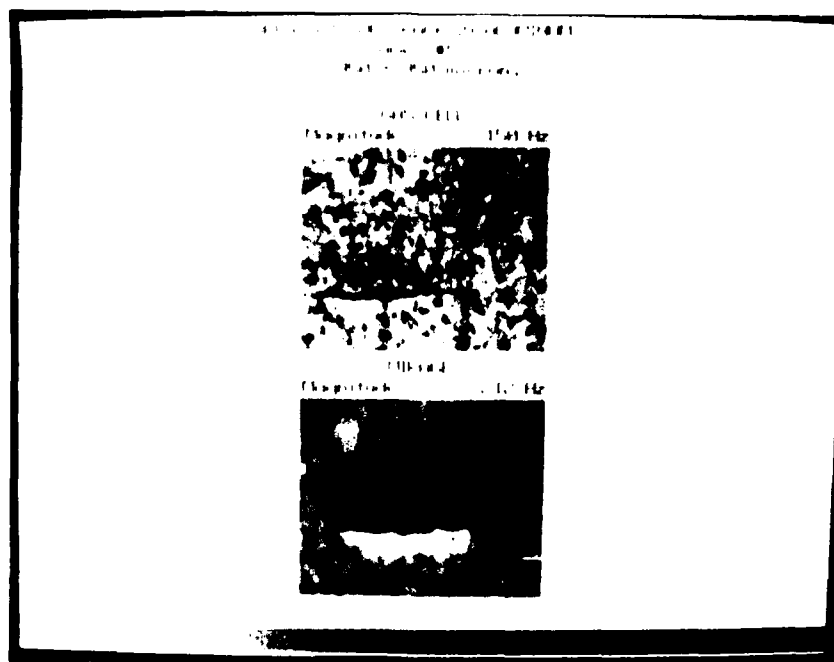
Figure 10. Fatigue crack in aluminum: Gas Cell vs mirage detection.

Figure 11 is an image of an area at the trailing edge of a first stage turbine blade for the TF30 engine. Two cooling air holes running through the blade beneath the surface are visible as two broad white horizontal bands in the image. This particular image was created by merging 3 smaller images side by side and using an image processing technique known as convolution to reduce some experimental artifacts that occurred at high spatial frequencies. Not shown for lack of space are images showing a feature in an area of the blade that had a previous radiographic indication.
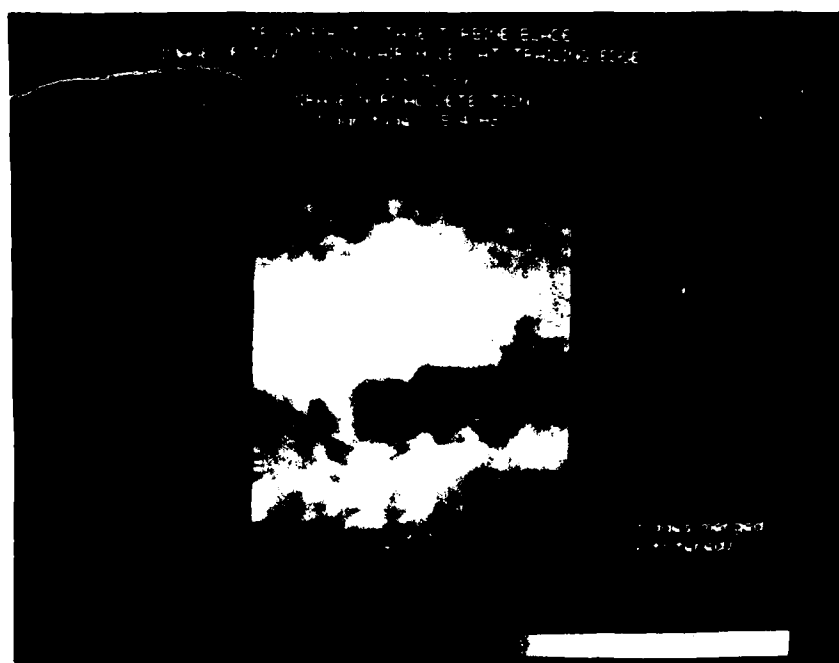
Figure 11. Image of subsurface turbine blade air cooling holes.

Figure 12 is an "ion acoustic" image of a polished single crystal slab of zirconia made by using a chopped ion beam and piezoelectric detection (4). Ion implantation can improve wear resistance as well as offering better corrosion, fatigue and hardness properties on a wide variety of materials. Nitrogen implantation such as used in this case can increase the hardness of the zirconia (16). The white patch in the upper left of the in phase ion acoustic image in Figure 12 corresponds to an implanted area. Only a very faint yellowish area was visible to the eye by using a glancing reflection. The transparency of the crystal rendered our usual photoacoustics with green Argon light ineffective because there was so little absorption. We see the best potential of ion acoustics in "on-line" control of ion implantation on complex or inhomogeneous surfaces. Similarly, we feel that the photoacoustic effect can also be useful for control of materials processing via lasers.
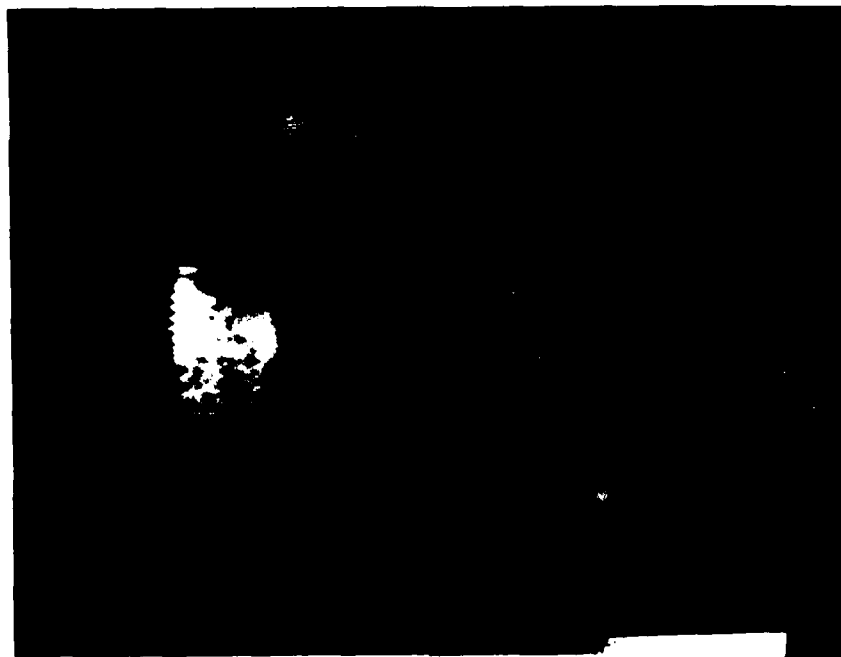
Figure 12. Ion acoustic image of a nitrogen implanted area in a zirconia crystal.

CONCLUSION

Photoacoustic microscopy offers another way of "looking" at surfaces which can be useful for characterization and near surface flaw detection down to about 2mm. At present it is relatively immature and slow. The effect holds promise for on-line control of laser and ion beam processing.

ACKNOWLEDGEMENTS

REFERENCES

1. A. Rosencwaig, "History of Photoacoustics", Chapter 2 in Photoacoustics and Photoacoustic Spectroscopy, John Wiley and Sons, New York, New York (1980).

2. Y. H. Wong, R. L. Thomas, and G. F. Hawkins, "Surface and Subsurtace Structure of Solids by Laser Photoacoustic Spectroscopy", Appl. Phys. Lett. 32, 538 (1978).

3. E. S. Cargill, III, "Electron-acoustic Microscopy", Phys. Today, 34, 27 (1981).

4. K. O. Legg and D. N. Rose, "Scanning Ion Acoustic Microscopy for Analyzing Ceramic Surfaces", Proceedings of the 8th Annual Conference on Composites and Advanced Ceramic Materials, 15-18 Jan 1984, Cocoa Beach, Florida, to be published.

5. J. A. Noonan and D. M. Munroe, "What is Photoacoustic Spectroscopy", Optical Spectra, February (1979).

6. T. A. Moore, R. Tom, E. P. O'Harra, D. M. Anjo, and D. Benin, "In Vivo and In Vitvo Studies of Biological Materials by Photoacoustic Techniques", presented at AAAS meeting, 26-31 May 1983, Detroit, Michigan.

7. R. L. Thomas, L. D. Favro, P. K. Kuo, and D. N. Rose, "Scanning Photoacoustic Microscopy of Aluminum with Aluminum Oxide, Roughness Standards and Rubber", US Army Tank-Automotive Command Research and Development Center TR 12668 (1982).

8. J. C. Murphy and L. C. Aamodt, "Signal Enhancement in Photothermal Imaging Produced by Three Dimensional Heat Flow", Appl. Phys. Lett. 39, 519 (1981); L. C. Aamodt and J. C. Murphy, "Photothermal Measurements Using a Localized Excitation Source", J. Appl. Phys. 52, 4903 (1981).

9. A. C. Boccara, D. Fournier, and J. Badoz, "Thermo-optical Spectroscopy: Detection by the Mirage Effect", Appl. Phys. Lett., 36, 130 (1980); A. C. Boccara, D. Fournier, N. Jackson, and N. M. Amer, "Sensitive Photothermal Deflection Technique for Measuring Absorption in Optically Thin Media", Opt. Lett. 5, 377 (1980).

10. R. L. Thomas, L. D. Favro, K. R. Grice, L. J. Inglehart, P. K. Kuo, J. Lhota, and G. Busse, "Thermal Wave Imaging for Nondestructive Evaluation", p 586, Proceedings of the 1982 IEEE Ultrasonics Symposium, B. R. McAvoy, ed., IEEE Press, New York, New York (1982).

11. A. Rosencwaig, "High Resolution Thermal Wave Imaging" presented at AAAS meeting 26-31 May 1983, Detroit, Michigan.

12. G. Busse, "Photothermal Remote Nondestructive Material Inspection", presented at AAAS meeting 26-31 May 1983, Detroit, Michigan.

13. M. Luukkala, J. Jaarinen, and A. Lehto, "Photothermal Measurement of the Thickness of Diffusion Hardened Surface Layers in Steel", Proceedings of the 1983 IEEE Ultrasonics Symposium, IEEE Press, New York, New York, to be published.

14. P. K. Kuo, L. J. Inglehart, L. D. Favro, and R. L. Thomas, "Experimental and Theoretical Characterization of Near Surface Cracks in Solids by Photoacoustic Microscopy", p 837, Proceedings of the 1981 IEEE Ultrasonics Symposium, B. R. McAvoy, ed., IEEE Press, New York, New York (1981).

15. P. K. Kuo and L. D. Favro, "A Simplified Approach to Computations of Photoacoustic Signals in Gas Filled Cells", Appl. Phys. Lett. 40, 1012 (1982).

16. J. K. Cochran, K. O. Legg, and G. R. Baldau, "Microhardness of N-Implanted Yttria Stabilized ZrO_2", presented at the Nineteenth University Conference on Ceramic Science, 8-10 November 1982, North Carolina State University, Raleigh, North Carolina.

DIGITAL BEAMFORMING ARRAY(U)

*JOHN F. ROSE, MR.
JESTER M. LOOMIS, DR.
ARMY MISSILE LABORATORY
REDSTONE ARSENAL, AL 35898

I. INTRODUCTION

The concept of forming receive antenna beams, in a computer through the use of the digital representation of the elements amplitude and phase has been pursued for many years [1, 2, 3]. This technique which has been called array signal processing or digital beamforming (DBF) has several advantages when compared to conventional beamforming techniques. One advantage that DBF brings to array technology and ultimately radar system performance is the flexible ease of forming multiple beams through the parallel use of digital algorithms such as FFT's rather than replicating different microwave feeds. Multiple beams give the radar designer another degree of freedom in that he can trade time and power to increase surveillance coverage, or to spend more time at a given point in space. Longer dwell times will enhance clutter processing and may ultimately, through the exploitation of narrow doppler filters, provide a means of target classification. In addition, having a digital representation of the coherent phase and amplitude at each element gives the potential to perform adaptive beamforming, nulling, and even extended resolution techniques such as maximum entropy.

Many new developments are occurring in electronic components and these advances now make it possible to consider DBF for military applications. For example, real time data processing limitations have been an impediment to this concept since its inception, but recent increases in computational ability developed through the DOD VHSIC program have brought processor speeds up to levels necessary to support many radar requirements. In addition, the RF and IF components being developed in microwave monolithic circuits offer the very real possibility of reduced size, cost, and power consumption necessary to afford and package a receive/ downconversion module behind each element. Current state of the art, as well as projected advances in microcircuit electronics, will add improvements in the speed, reliability and repeatability of an element module.

157

These technologies, coupled with component advances in fiber optics, subarraying techniques, etc., begin to make DBF very attractive for military use.

In a conventional radar antenna the performance parameters are a function only of the microwave components; whereas in a DBF array not only the microwave components, but also downconversion and digital components affect pattern performance. At the same time, radar measurement parameters such as range and radial velocity, which heretofore have been independent of the array, now are affected by the array design and element module component selection. In order to examine these effects, a prototype DBF antenna and radar were designed and built. The remainder of the paper summarizes the unique design considerations associated with the DBF system and presents the results of performance testing.

II. DESCRIPTION OF DBF PROTOTYPE HARDWARE

A block diagram of the eight-element antenna system is shown in Figure 1. Behind each element of the array is a receiver to convert the received RF signal into a format suitable for digitizing. The receive element of the array is an H-plane sectorial horn, which was selected because of its simplicity, and the desire to obtain some directivity in the azimuth plane while maintaining a wide element pattern in the elevation scan plane. The inter-element spacing was chosen to be 0.6" in order to eliminate grating lobe effects. This spacing also could accommodate a commerically available power divider that permitted direct connection of the antenna, RF mixers, and local oscillator (LO) feeds. Such a packaging technique eliminated the need for high tolerance RF interconnecting cables.

In order to establish the pattern characteristics of the array without receiver circuitry it was necessary to measure the individual element reflections and the mutual coupling between elements. Following a procedure described in [4], the reflection coefficient was measured for each element in the array environment. The coupling between elements was determined by exciting one element and measuring the received signal at each of the other elements with all other elements properly terminated. Typically about 0.1 dB of amplitude mismatch and 2.5° of phase deviation were attributed to the reflection and coupling effects. The effect of these variations is shown in Figure 2, where a no-error pattern (shown dotted) is compared to the pattern with errors. The errors result in virtually no change in the main beam but cause a 3 to 5 dB increase in sidelobes. This errored pattern now becomes the baseline pattern from which the receiver error effects can be quantified.
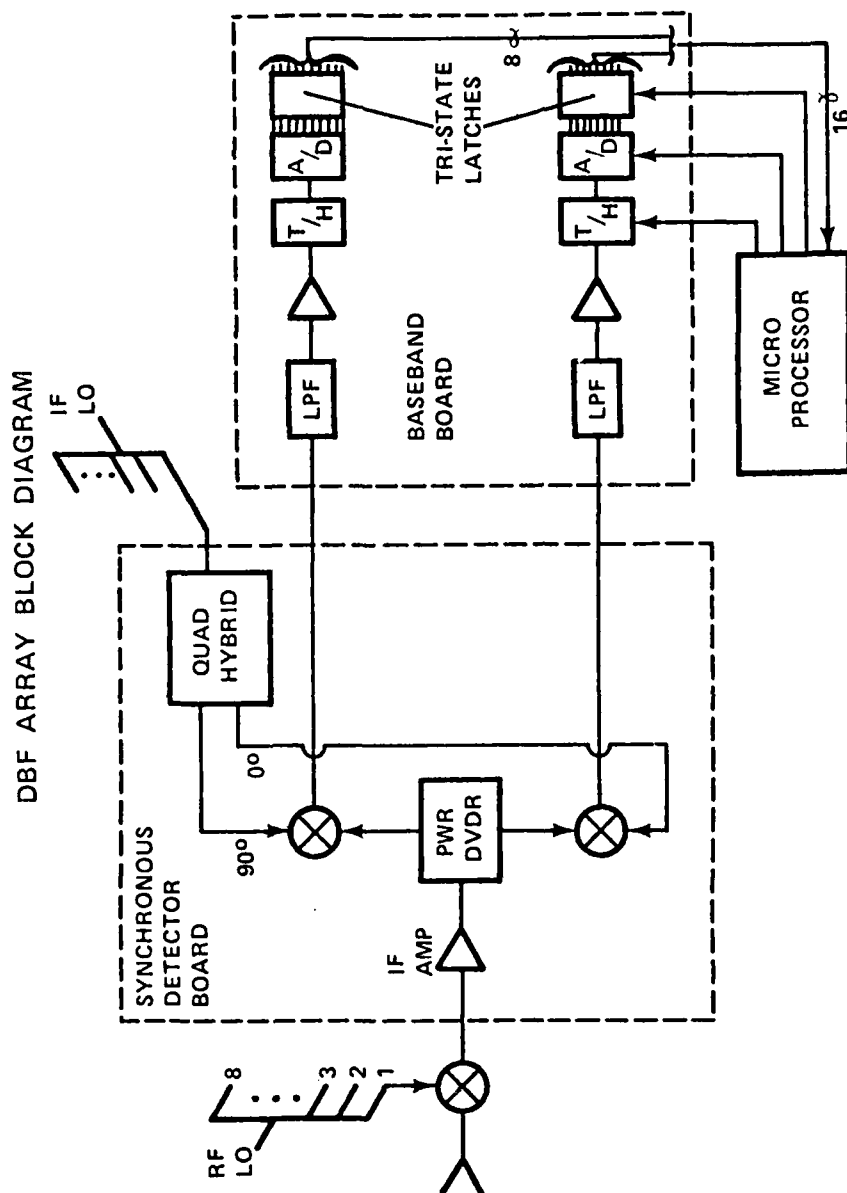
Figure 1. Antenna system block diagram. This block diagram shows one element receive chain of the eight element array.
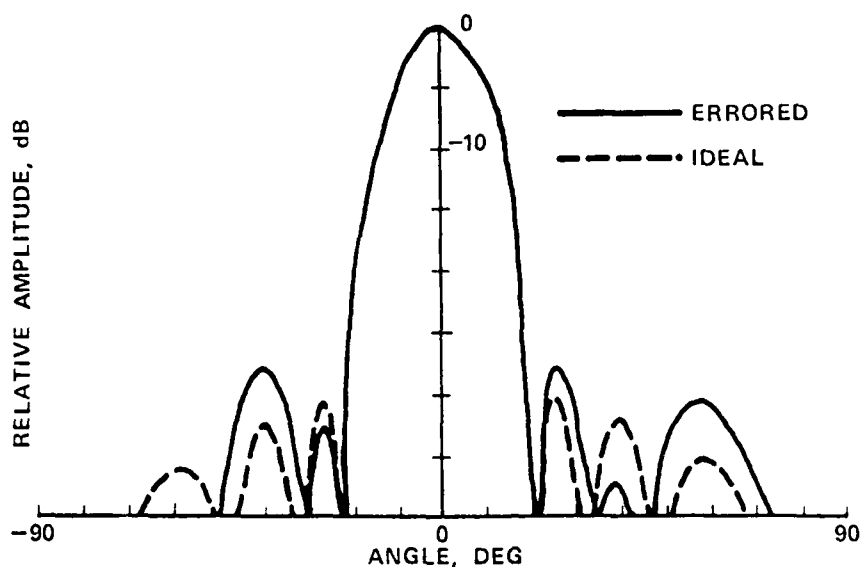
Figure 2.   Ideal pattern (dotted) and computed pattern (solid)
             when reflection and coupling errors are introduced.

        Return to Figure 1 where the receiver circuitry behind each element
is shown. The X-band signal coming from each array element is first down-
converted to an IF frequency of 30 MHz. The signal is then amplified and
mixed down to base band in a synchronous detection process. The output of
this stage is an in-phase (I) and quadrature (Q) representation of the
received signal. The critical parameter which must be considered in the
design and fabrication of this unit is the thru phase and gain of each
path which must be matched to preserve the relative amplitude and phase
of the received signals. In addition, the dynamic range of the system
must be considered so that none of the stages in the receiver saturate and
create a nonlinear phase and amplitude response. Table 1 shows each
element's relative phase for the total downconversion process. This data
was taken by combining the RF mixers, LO distribution power dividers, and
synchronous detector boards and measuring the output phase differences
between I channels. The phase of the Q channel was determined by first
measuring the difference between the I and Q phase on each individual
board and then adding this difference to the measured I channel phase.
The standard deviation of phase error in the I and Q channels was computed
to be 3.3°.

Table 1. Relative phase of downconversion
process from RF mixers through
synchronous detection

Element	Relative phase	
	I	Q
1	0	89
2	7.6	97.6
3	7.1	96.1
4	3.7	96.7
5	0.7	88.7
6	3.4	93.4
7	8.0	95.0
8	8.2	95.2

The final signal conditioning in the elemental receive channels is baseband amplification. In this stage, the I and Q signals are amplified to bring their levels up into the dynamic range of the A/D converter. In addition to acting as an amplifier, this section also performs necessary filtering functions and amplitude gain control. Following the amplifiers, a track-and-hold (T/H) circuit is used to present the A/D converter a fixed voltage for digitizing. A fixed level is necessary to reduce errors that occur if signal levels change during the conversion time.

The selection of the 8 bit Analog Devices HAS-0802 A/D converter was based on obtaining an off-the-shelf A/D with relatively simple interface requirements, having a moderate dynamic range (7 bits + sign = 42 dB) and a conversion time fast enough to sample a radar waveform and provide less than a 300 m range cell.

There are two types of A/D errors that affect array performance: quantization and linearity. Quantization errors are inherent in an A/D because the converter approximates an analog signal with a discrete digital word. The quantization error is determined by calculating the variance (power) of the distribution function over an interval of 1 LSB and is equal to $q/\sqrt{12}$ where q is the LSB. The second source of errors can be broadly classed into a term called linearity errors. Linearity is a measure of the variation in the difference in input voltage necessary to change the output of the A/D. Or, in other words, linearity is the variation of the LSB between different digital states. There are two types of linearity errors, the first is a random error, bit state to bit state, its effect is similar to a raised noise floor. The second type of linearity error is a systematic error in the adjacent LSBs. For example, many A/D's exhibit a variation in LSBs which is quadratic such that the LSB at the higher levels is larger than the LSB at the low conversion

levels. The result of systematic linearity errors is the creation of harmonic responses in a frequency or spatial transform process.

Throughout this section only the phase errors have been discussed at any length. Generally, these errors are not only much more critical to array performance, but also are more difficult to control than amplitude errors. In the prototype system a gain control was put in the baseband amplifiers, in order to balance the gain variations in the receive circuitry. The gain control was placed ahead of the track-and-hold and A/D convertor because the errors which occur in these components are very small and have little effect on the array's performance.

III. ANTENNA PATTERN DATA

Upon assembly of the complete array, antenna pattern data was taken to establish performance bounds and relate component error effects to pattern characteristics. In a DBF array several features require that a somewhat unconventional approach be taken toward both data gathering and data reduction. For example, a conventional antenna pattern is taken by recording, typically on chart paper, the output of a single beam port as a function of angle; however, in the DBF array all element outputs are sampled as a function of angle and the actual pattern is formed afterward. The pattern is not unique in that the same data can be weighted in many ways, depending upon the desired purpose. This feature was used as a data reduction and analysis tool in order to form different types of beams to accentuate different error effects.

A block diagram of the equipment and interfaces used to take antenna patterns is shown in Figure 3. The transmit signal is derived from mixing the X-band LO and a stable 30.010 MHz source. This combined signal, after appropriate filtering and amplification, was transmitted from a remote source, received, and downconverted to 0.010 MHz or 10 kHz at each element. The offset is necessary for the signal to pass through the baseband amplifiers. The actual transmit horn was mounted on a tower 31 feet away from the receive array. In order to calibrate the array, the array was placed normal to the transmit source and gain adjustments were made by equalizing the I and Q rms voltage in each element. This technique will account for all amplitude imbalances caused by the downconversion and final amplification process, as well as any amplitude imbalances in the antenna aperture due to mutual coupling effects at broadside. Coupling differences due to illumination functions variation and scanning will not be accounted for. This method works well under the controlled conditions of an antenna range where the source can be placed in the far field and multipath can be eliminated; however, in radar applications such controlled conditions cannot generally be enforced and some other calibration technique must be developed. Radar calibration will be discussed later in the paper.

ANTENNA PATTERN
EXPERIMENT CONFIGURATION
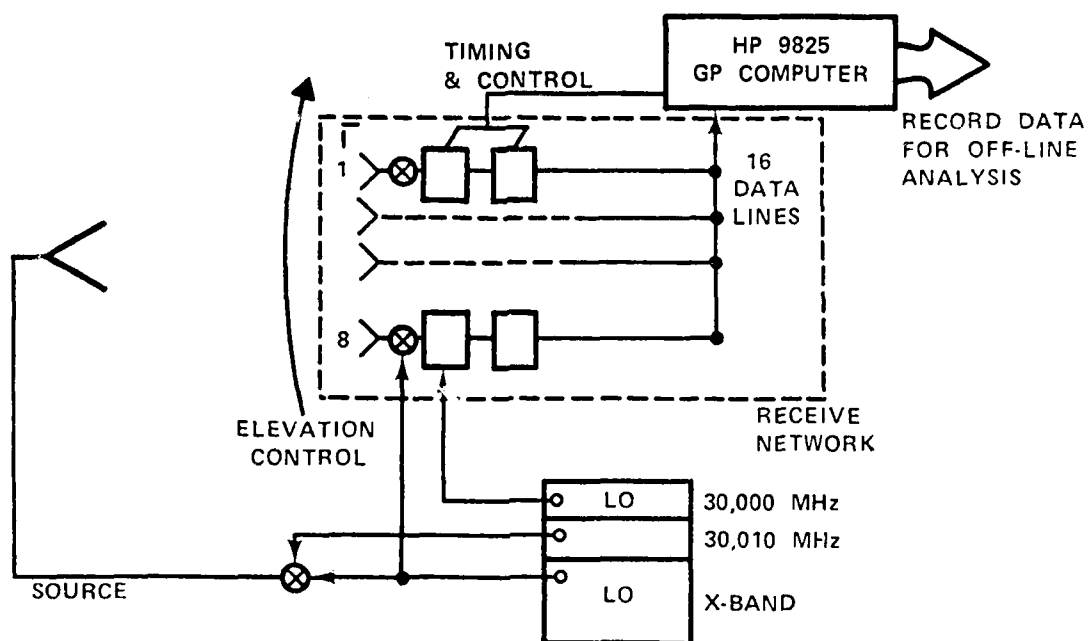
Figure 3. Block diagram of pattern recording equipment and
receiver test network.

Upon calibration of the array the position and element output data
were recorded for future data analysis. The timing control and data
recording is done by an HP 9825 microcomputer. Since the data was stored
on tape in the HP 9825 it was very easy to use the 9825 for data reduction
also. Two general classes of programs were written to analyze the data.
The first type was a group of programs which examine the elemental data to
see if it is valid or to specifically examine element-to-element signal
characteristics. The second type of programs were developed to form pat-
terns with combined element data, just as a DBF spatial processor would
do.

The most useful analysis performed on the individual element data was
the computation of phase and amplitude errors. The phase errors are com-
puted by applying a least squares fit to the linear phase front across the

163

array. The difference between the least squares line and the calculated phase at an element is defined to be the error at that element. This is not an absolute error because the least squares line does not represent the actual phase slope, only an approximation with respect to the data. The amplitude errors are found by computing the average magnitude across the array and comparing each element's magnitude to this average. Ideally, the amplitude errors are very small; however, the amplitudes were balanced with respect to broadside coupling so that as the antenna is rotated and coupling changes a residual error results. This residual error was calcu-lated to be less than 0.25 dB over a plus or minus 60 deg. scan range.

The effect of the component errors is a departure of the realizable antenna pattern from the ideal. This departure results in a loss of gain, an error in pointing direction, and an increase in sidelobe level. For small amplitude and phase errors the effect on gain and pointing is very difficult to measure and quantify; whereas, the error effect on sidelobes is usually dramatic, especially if the sidelobes are designed to be low initially. Consequently, in order to assess the errors, low sidelobe illumination functions were applied to the data to accentuate the error producing effects. Figure 4 depicts the pattern obtained when the data is weighted with a 30 dB Chebyshev illumination function.

Using a method described in [5], the pattern results indicate that the total component errors in the array are 3.6^o. In order to establish the mutual coupling error level the unerrored pattern of Figure 2 can be compared to the pattern derived from only the mutual coupling data also shown in Figure 2. Based on these patterns the rms error due to mutual coupling is about 1.1^o. Since these errors are independent of the inser-tion phase and I and Q sampling variations, the errors due to the insertion and sampling variations, e_1, can be computed to be

$$e_1 = \sqrt{e^2 - e^2_c} \quad = \quad \sqrt{3.6^2 - 1.1^2} \quad = \quad 3.4^o \, , \tag{1}$$

where e is the total error and e_c represents the mutual coupling errors.

The rapid fluctuations in the sidelobes of the pattern in Figure 4 are not typical of a conventional array, whose patterns look more like the pat-tern of Figure 2. The variations in the sidelobes are caused by a dif-ferent error set being applied to the array as the received waveform is sampled in time. In order to differentiate the I and Q sampling error from the time invariant insertion phase errors, the design level is taken to be the average of the fluctuations over a relatively constant portion of the pattern. The angular region between 45 and 65° in Figure 4 was

164
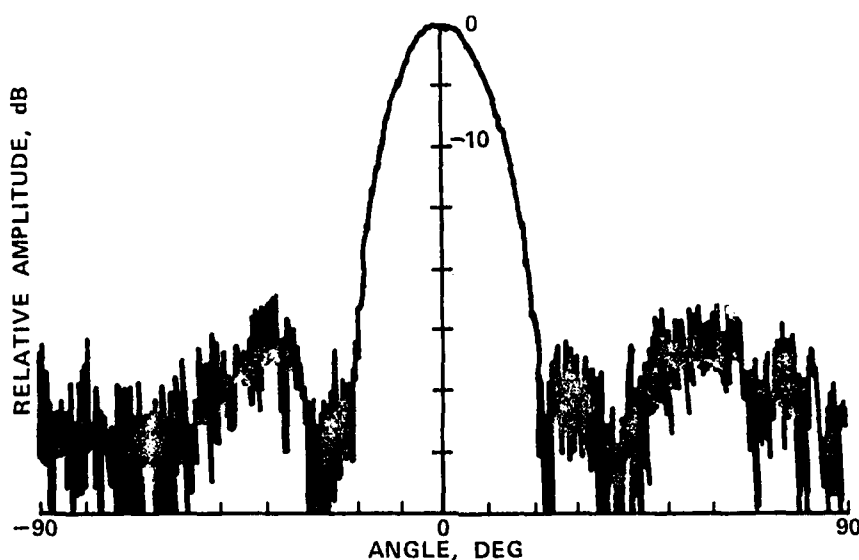
Figure 4. Pattern resulting from DBF array when elements data is weighted by a 30 dB Chebyshev illumination.

selected for the calculations. Following the same procedure used earlier results in a measured rms I and Q sampling error of $1.7°$. Subtracting this error variance from e_1 as found in Equation (1) results in a fixed *RMS insertion phase error of $2.9°$*. This compares to an error of $3.3°$ measured on the actual components (see Table 2). Based on the component data, the expected RMS error due to I and Q sampling variations for a 10 kHz baseband frequency is $2.0°$ RMS which is very close to the $1.7°$ error computed from the measured data. Table 2 summarizes the error analysis by comparing the measured error of each subsystem to the error calculated from the antenna patterns.

Table 2. RMS phase error summary

Error Source	Measured error, deg	Calculated error, deg
a. Refection and Coupling	2.6	1.1
b. Element to element down conversion	3.3	2.9
c. I to Q due to down conversion	1.3 ⎱ 2.0	1.7
d. I to Q due to filters	1.5 ⎰	

IV. RADAR OPERATION

Upon completion of the pattern tests a prototype radar was built around the receive array. Basically this consisted of adding a transmit antenna, transmitter, modulator, timing and control functions to the DBF array. During this transformation, the IF and baseband amplifier gains were changed to help reduce line voltage contamination of the radar signals and the IF amplifier/synchronous detector section was repackaged to improve their reliability. In addition, a pilot pulse calibration mechanism was added to the receive network. The calibration network consists of a separate oscillator offset 10 kHz from the intermediate frequency (30 MHz) LO, an additional 8 way power divider and mixer, and couplers behind each element. A radar block diagram is shown in Figure 5.

The radar covers a fixed 12^0 sector in azimuth and forms beams in elevation over a $+90^0$ sector. Target illumination is provided by a TWT transmitter radiating through an 11 dB gain flood illuminating horn. The TWT is pulsed for 2 μsec at a pulse repetition frequency (PRF) of nominally 30 kHz. These parameters, coupled with the receiver characteristics, result in a detection range of about 5 km on a 5 m^2 target after spatial and signal integration. As in the array case, the return signals are received by each element, down converted to base band, filtered, and finally converted to a digital word. The A/D conversion rate is 500 kHz, which corresponds to a sample every 2 μsec or every range cell. There is a separate line from each I and Q channel to the computer, making a total of 16 parallel lines from the eight element array. The clock rate is 5 MHz, broken up so two clock intervals are used for house-keeping and timing functions and eight clock intervals are used to transfer eight bits of data from each channel. This data rate is too fast to be accepted by a general purpose computer; therefore, it was necessary to build a memory buffer to preprocess the data for computations. The memory buffer is a 1024 word by 16 bit device. This results in 128 data samples (1024/8 elements) that can be used for range and doppler processing. The processor being used is a Digital Equipment PDP-11/34 with an AP120B array processor to perform the Fourier Transform operations. A symbolic diagram of the processor is shown in Figure 6. The processor is flexible in that it can process one pulse over 128 range cells, 128 pulses over one range cell, or any binary combination such as 2 X 64, 4 X 32, etc. in between. The control signals to the A/D converter were developed in order to sufficiently utilize the 128 available sample points. This is done by sampling only a group of range cells about a desired range, rather than continuously sampling between transmit pulses.
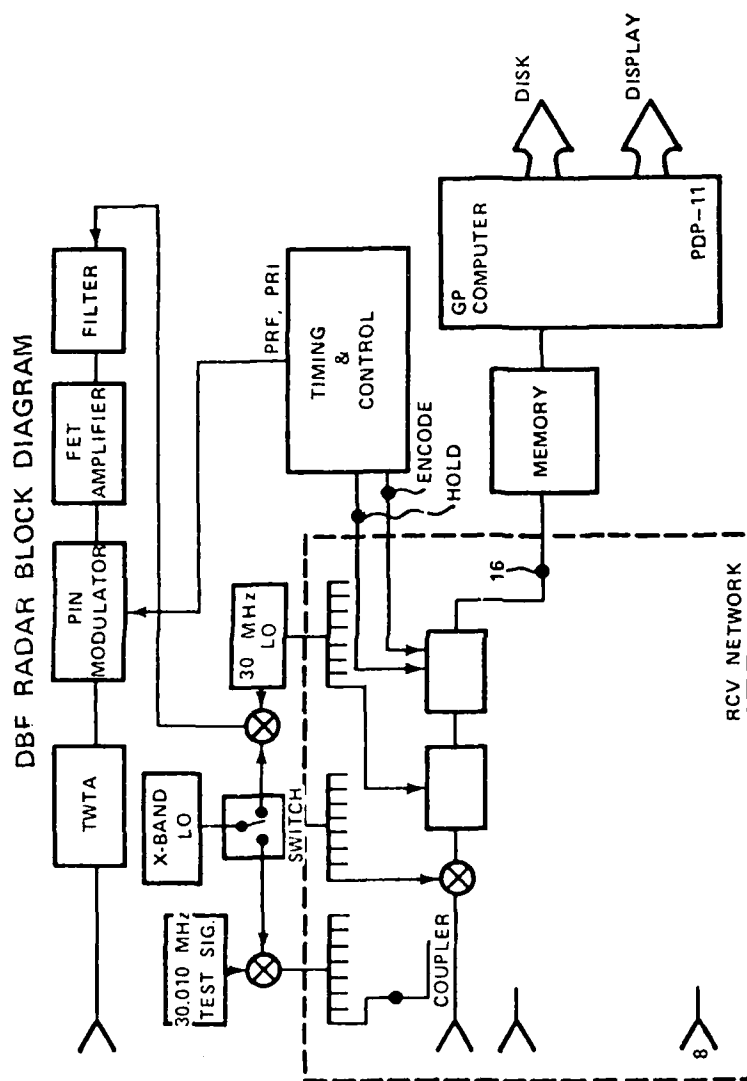
Figure 5. Block diagram of digital beamforming radar.
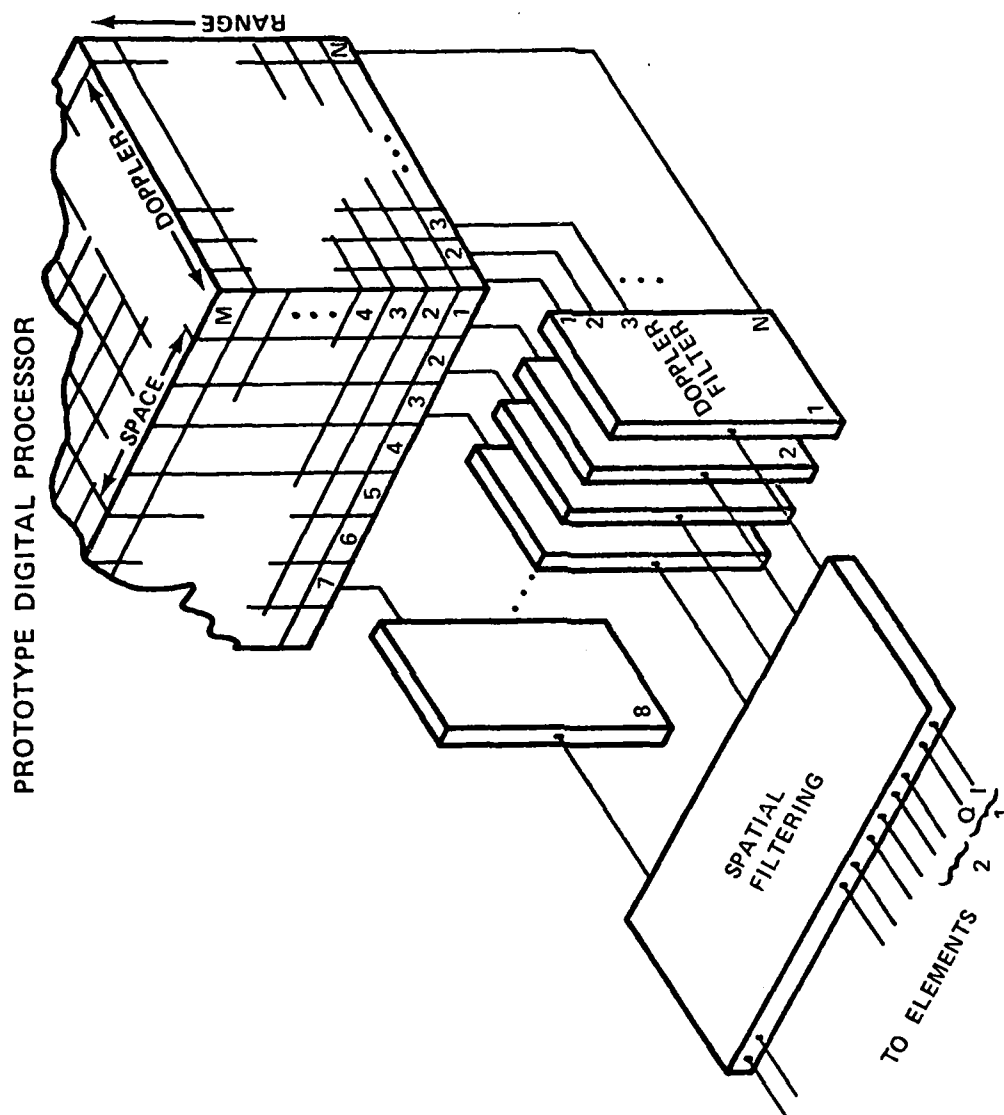
Figure 6. Block diagram of DBF radar signal processor.

For example, if only four range cells are sampled following each pulse, a 128/4 or 32 point doppler transform can be performed on each block of buffered data.

A pilot pulse injection scheme was used in the radar configuration for receiver calibration. This was necessary because the radiating source technique used during the antenna tests was not practical because of interference caused by surrounding objects. A pilot signal offset from 10 GHz by the 30 MHz LO and by an additional 10 kHz doppler was injected via the power divider and couplers into the front end of each receive channel. A Discrete Fourier Transform (DFT) was applied to the data from each channel. By examining the outputs of the DFT it is possible to determine the mismatches in the through-phase of the receivers and the individual I and Q channel errors. These errors are compensated for by applying a phase and amplitude adjustment to the illumination function weights in both the doppler and spatial filtering operations.

In order to evaluate the performance of the DBF radar a test target generator (TTG) was used rather than trying to work with targets of opportunity. The TTG is a transponder which returns the transmitted signal with a doppler offset frequency. It was usually placed about 1 km from the radar during tests. A test simply consisted of radiating toward the TTG and then receiving the transponded signal. In this way, the range, angle and doppler of the return can be carefully controlled.

Figure 7 shows a target response resulting from the radar tests. The 3-D plot has axes of doppler cell (0-31), beam position (0-7) and return amplitude in dB. This plot is a horizontal slice from the processor diagram depicted in Figure 6. The large amplitude signals on the beam axis at zero doppler are due to clutter. The target is located in beam 4 in the tenth doppler cell. It is the largest return other than those due to clutter. The amplitude in the non-target cells along the axes orthogonal to the target are substantially larger than those in the off axis region. This is due to the target showing up in the sidelobes of the beam and doppler cells of the radar. The spatial sidelobes are less than those along the doppler axis because a low sidelobe weighting function was used in space; whereas, uniform shading was used in doppler.
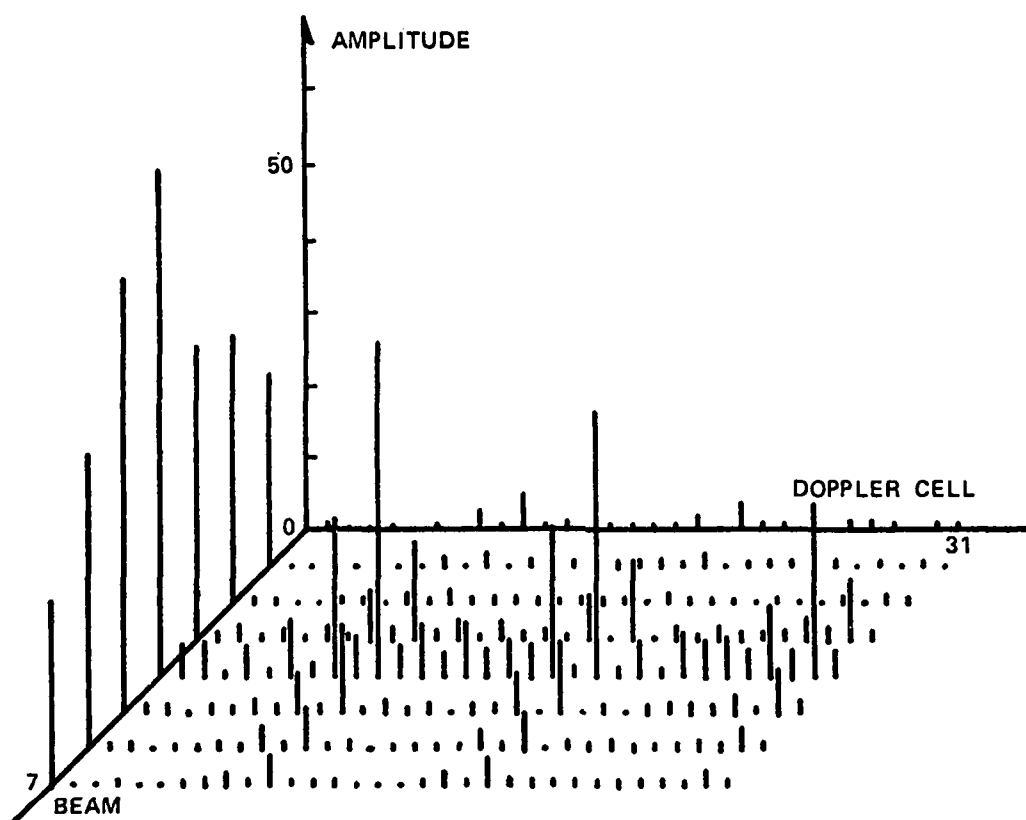
Figure 7.   Response of DBF radar.

V.   CONCLUSIONS

The work described herein has demonstrated the feasibility of using an antenna whose beams are formed digitally for radar applications. Such a radar antenna has flexibility in its shapes and characteristics. For example, radar beams with different sidelobe levels, difference beams, and special shaped beams such as cosecant squared beams can be formed simultaneously. Moreover, the quality of these beams can be made to be equal to or better than that obtained from conventional fixed microwave feeds.

The additional elemental hardware necessary for reception and digitization makes reproducibility more difficult; however, the ability to measure array errors in real time and the ease of adjusting the weighting coefficients ultimately offsets the increased error sources in the array. Two techniques were used to provide calibration signals for array measurements in the prototype system. The remote illumination method used in the antenna tests was very successful in the relatively pure environment of an antenna range. This scheme is not practical in a radar configuration because of interference by nearby objects. In this case a signal injection technique produced performance similar to that of the illumination method. Its principal limitation was the stability of the injection signal hardware.

The radar tests demonstrated that the receive hardware which now is replicated many more times than in a conventional array does not adversely impact the performance of the signal processing functions. Range and doppler processing was carried out at a performance level commensurate with that achieved with current radar practice.

The major technological difference between the prototype array described here and one that could be used in the field Army is in the number of elements and the signal bandwidth. A typical Army radar would require a factor of ten increase in the number of elements and about a five-fold increase in bandwidth. With the projected advances in the operational speed of digital hardware, it will be feasible to build an Army radar which uses digital beamforming in the future.

REFERENCES

1. P. Barton, "Digital Beam Forming for Radar IEE Proc.," Vol 127, Pt. F, No. 4, pp 266-277, August 1980.

2. W. Sander, "Experimental Phased-Array Radar ELRA: Antenna System IEE Proc.," Vol 127, Pt. F, No. 4, pp 285-289, August 1980.

3. D. E. N. Davies, "A Fast Electronically Scanned Radar Receiving System," Br. Inst. Radio Eng, J., Vol 21, pp 305-318, 1961.

4. A. A. Oliner and R. G. Malech, "Mutual Coupling in Infinite Scanning Arrays," of Microwave Scanning Antennas, Vol II, edited by R. C. Hansen, Academic Press, New York, 1966, Ch. 3.

5. J. L. Allen, "Phased Array Radar Studies," MIT, Lincoln Laboratory, Lexington, Mass., 1960-61.

HEMORRHAGIC FEVER WITH RENAL SYNDROME:
DISCOVERY OF A UNIQUE GROUP OF VIRUSES

Dr. Connie S. Schmaljohn, Mr. Sherman Hasty and Dr. Joel M. Dalrymple
Department of Viral Biology, Virology Division
United States Army Medical Research Institute
of Infectious Diseases
Fort Detrick, Frederick, MD 21701

Korean hemorrhagic fever has been recognized as a military health problem since it was first reported in U.S. troops during the Korean conflict, when approximately 3,000 cases occurred, 10% of which were fatal. This disease remains a threat to troops stationed in Korea and has resulted in numerous illnesses and deaths in the last 3 years. A large group of diseases which are clinically similar to Korean hemorrhagic fever are endemic throughout China, Russia, Japan, Europe and Scandinavia. The World Health Organization has collectively termed these diseases "hemorrhagic fever with renal syndrome" (1). A viral etiology for these diseases was suggested in 1978, when a virus was isolated from Korean striped field mice,(Apodemus agrarius corea), and was found to react with antibodies in the blood of patients with Korean hemorrhagic fever (2). The virus was named Hantaan after a river in Korea paralleling the demilitarized zone near which it was isolated. Adaptation of Hantaan virus for growth in cell culture significantly contributed toward charaterization of this disease causing agent (3).

Basic molecular and antigenic characterization of Hantaan virus revealed that it was similar to the Bunyaviridae family of animal viruses and possessed a three segmented, single-stranded RNA genome contained in three ribonuclease sensitive nucleocapsids, surrounded by an envelope with two glycoproteins and a virion associated polymerase (4,5,6). Two-dimensional electrophoresis of ribonuclease T1 resistant oligonucleotides (fingerprinting) (7) indicated that the large (L), medium (M), and small (S) RNA segments of Hantaan virus are different from each other and from host cell RNAs (figure 1).

No serological relationship between Hantaan virus and viruses within the four existing genera of Bunyaviridae has been documented. However, it was recently determined that Bunyaviridae can be assigned to a genus by molecular analysis of their 3´-terminal nucleotide sequence which has been reported to be conserved for viruses within a genus but different from viruses in other genera (8).

Sequence analysis of the Hantaan virion RNA 3´ termini demonstrated that like the Bunyaviridae, the three RNA segments of Hantaan had the same 3´-terminal sequence, but that this sequence was unlike those of all other genera of Bunyaviridae (figure 2 and table 1). These data led us to propose that a fifth genus of Bunyaviridae, to be known as the Hantavirus genus is required to accomodate Hantaan-like viruses (5,6).
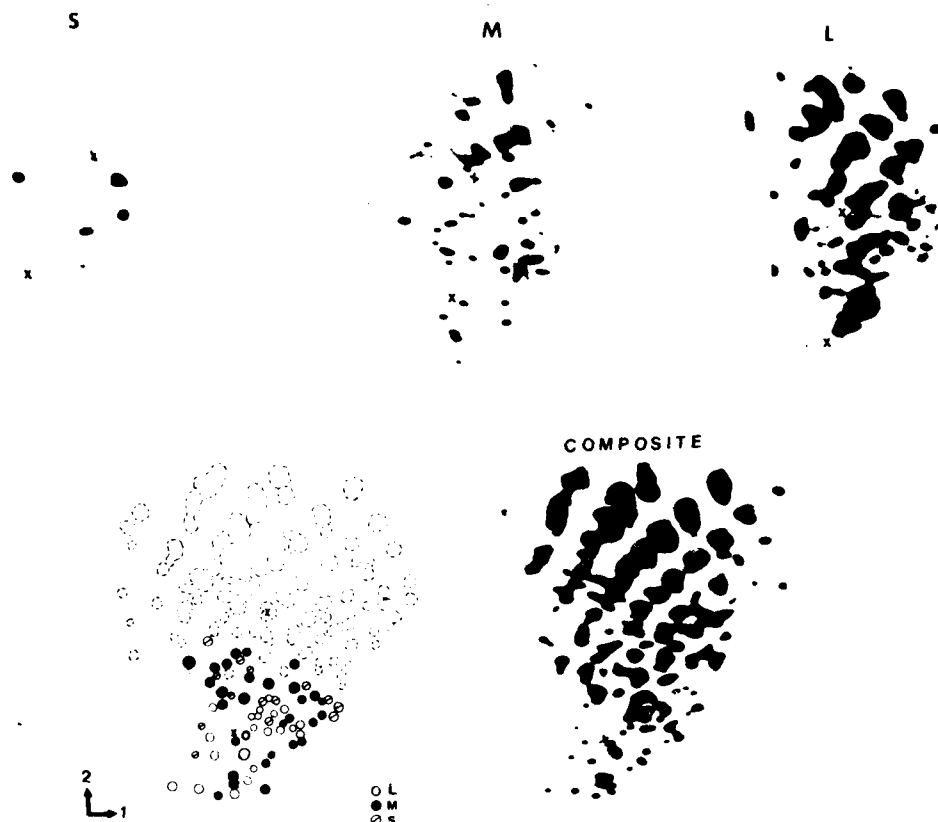
Figure 1. The three RNA segments of Hantaan virions were examined by two dimensional oligonucleotide mapping following digestion with ribonuclease Tl (fingerprinting) (5). Radiolabeled small (S), medium (M) and large (L) segments were examined individually and as a composite of total virion RNA. The schematic diagram indicates the location of the oligonucleotides from each segment within the composite map, and illustrates that each segment is unique.
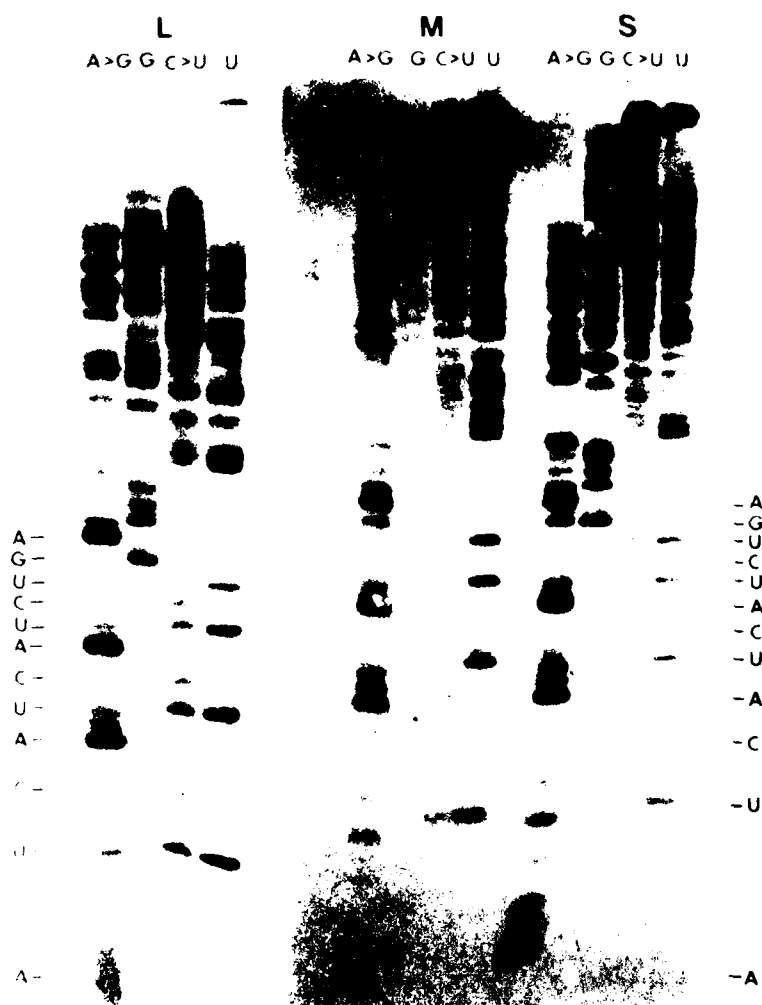
Figure 2. The 3´ termini of Hantaan virion L, M, and S RNAs were radiolabeled with PCP^{32} as previously described (5) and were sequenced using the chemical method of Peattie (9). 20% polyacrylamide:6M urea gels were used to resolve framents of RNA and were electrophoresed at 1500V for 2 hr.

Table 1

3´ NUCLEOTIDE SEQUENCES OF BUNYAVRIDAE

GENUS

BUNYAVIRUS:	3´	UCAUCACAUGA
UUKUVIRUS:		UGUGUUUCU
PHLEBOVIRUS:		UGUGUUUCG
NAIROVIRUS:		AGAGUUUCU
HANTAVIRUS:		AUCAUCAUCUG

Viruses serologically related to Hantaan have recently been isolated from rodents both in endemic disease areas, as well as in countries (such as the United States) where no associated disease is recognized. In order to gain a better understanding of the antigenic and molecular properties of this group of viruses, and to test our hypothesis that Hantaan and related agents should be classified in a new genus in the Bunyaviridae, we selected eight virus isolates for detailed comparison.

The viruses utilized in this study represented diverse geographical areas and all known hosts and are listed in Table 2. Virus isolates from endemic disease areas were represented by three Korean isolates, a Japanese isolate and a Finnish isolate. Three virus isolates from the United States represented viruses not known to cause human disease. The currently known host-range of this group of viruses (rats, mice, humans and voles) as well as three clinically distinguishable diseases (Korean hemorrhagic fever, epidemic hemorrhagic fever, and nephropathia epidemica) were also represented by these isolates. Virus propagation in cell culture has been described (4). None of the isolates produced detectable cytopathology, and growth characteristics of the virus strains were quite similar. Maximum virus infectivity titers of approximately 10^7 pfu/ml were achieved eight to twelve days post infection, and plaque assays were performed as previously described (4).

Table 2
VIRUS ISOLATES RELATED TO HEMORRHAGIC FEVER WITH RENAL SYNDROME

VIRUS	HOST	LOCATION	DISEASE
HANTAAN[1]	FIELD MOUSE	KOREA	?
LEE[2]	HUMAN	KOREA	KHF
URBAN RAT[3]	RAT	KOREA	?
TCHOUPITOULAS[4]	WHARF RAT	NEW ORLEANS	?
GIRARD POINT[5]	WHARF RAT	PHILADELPHIA	?
PROSPECT HILL[6]	VOLE	FREDERICK, MD	?
SAPPORO RAT[7]	LAB RAT	JAPAN	EHF
PUUMALA[8]	VOLE/HUMAN	FINLAND	NE

Viruses were isolated by: [1,2,3] H.W. Lee et al., Seoul, Korea; [4] T. Tsai et al., CDC Atlanta; [5] J. LeDuc, USAMRIID; [6] C. Gajdusek et al. NIH; [7] Kitamura et al., Tokyo, Japan; [8] J. Dalrymple et al., USAMRIID.

All isolates exhibited antigenic cross-reactivity by immunofluorescent staining of infected cells. The more specific techniques of radioimmune assay (RIA) and plaque reduction neutralization tests (PRNT) were employed therefore, to differentiate antigenic groups. Viruses were compared using antibody produced by experimental infection of Wistar rats with each of the eight virus isolates, and infected cell culture lysate antigens. The assay procedure has been described (10). Antigenic relationships of the viruses by this method appeared to be closely correlated with the type of host from which the virus was isolated (Table 3).

Comparison of plaque-reduction neutralization titers of the isolates also revealed host-range antigenic similarities (table 4). With this test, however, all viruses could be differentiated from one-another which clearly indicated that no two isolates were antigenically identical. Collectively, these data provide serological support for the concept that the Hantavirus genus is composed of a number of related but unique viruses.

To determine whether all isolates could also be differentiated molecularly, we compared the composite oligonucleotide maps of each virus isolate. The results of this comparison indicated that all viruses were genetically different from each other and confirmed the serological finding that no two isolates were the same. As an example of the genetic heterogeneity of these viruses, the oligonucleotide maps of the two American wharf rat isolates are displayed in figure 3.
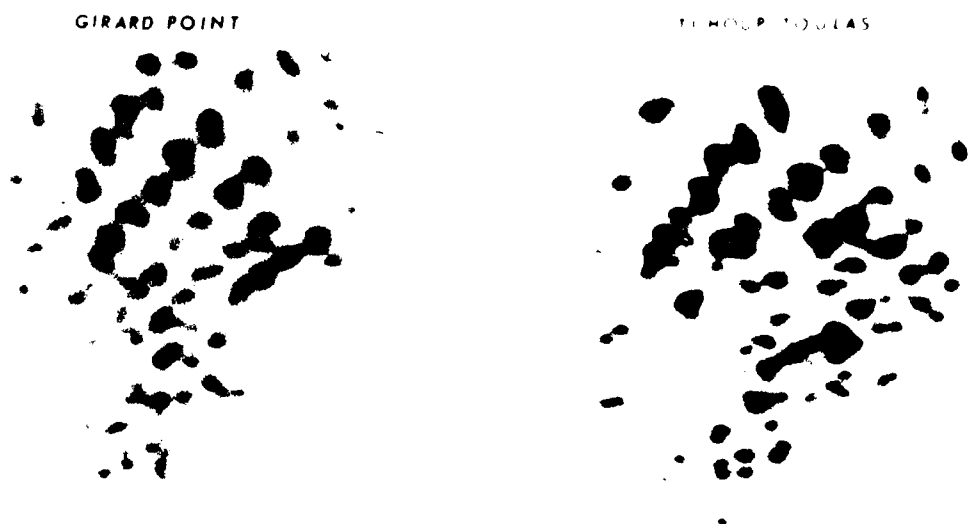
Figure 5. Comparison of the oligonucleotide maps of wharf rat isolates from New Orleans (Tchoupitoulas) and Philadelphia (Girard Point) illustrate the genetic diversity among the Hantaviruses.
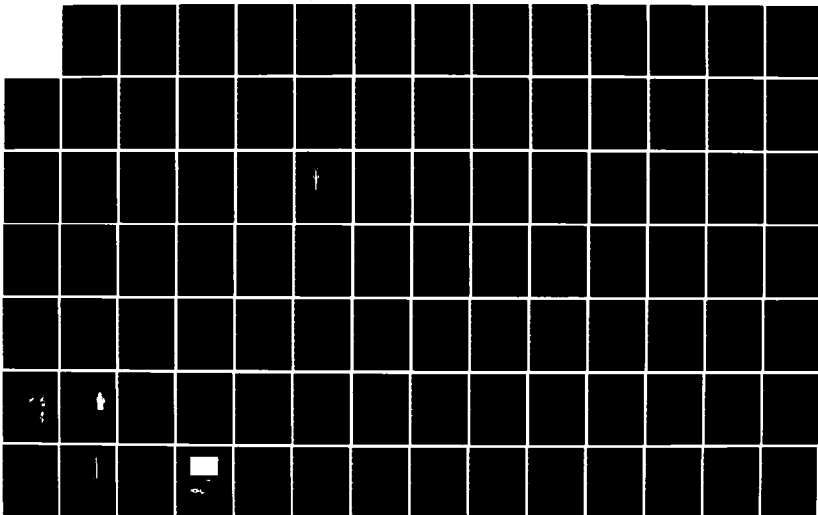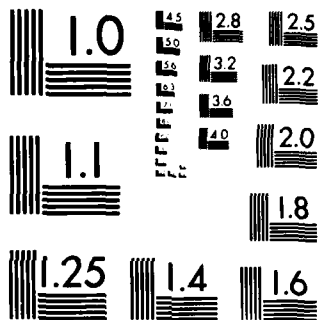
MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

Table 3

ANTIGENIC RELATIONSHIPS OF HFRS AGENTS
(RADIOIMMUNE ASSAY OF RAT SERA)

Serum Pools	Infected cell lysate								
	Hantaan 76-118	Lee	Urban Rat	Sap. Rat	Tchou-pitou.	Gir. Point	Pros. Hill	Puum-ala	Con-trol
Hantaan (76118)	100	30	30	30	30	30	10	3	.03
Lee (human)	100	100	30	30	30	30	10	10	.03
Urban Rat	100	10	100	100	100	100	10	10	.03
Sapporo Rat	30	30	100	100	100	100	3	10	.03
Tchoup-itoulas	100	30	100	100	100	100	10	10	.01
Girard Point	100	100	100	100	100	100	10	10	.03
Prospect Hill	3	3	3	10	3	3	100	10	1.0
Puumala (NE)	30	10	10	10	10	10	100	100	1.0
Control	0	0	0	0	0	0	0	0	0

For convenience, values have been expressed as % of the highest titer obtained with each antiserum (i.e. homologous = 100%).

Table 4   PLAQUE REDUCTION NEUTRALIZATION OF HFRS AGENTS

Serum Pools	Infected cell lysate							
	Hantaan 76-118	Lee	Urban Rat	Sap. Rat	Tchou-pitou.	Gir. Point	Pros. Hill	Puum-ala
Hantaan (76118)	4000	4000	-	40	160	nd	-	-
Lee (human)	4000	2000	20	200	160	nd	-	-
Urban Rat	2000	80	4000	8000	16,000	nd	160	-
Sapporo Rat	2000	20	2000	2000	2000	nd	200	-
Tchoup-itoulas	2000	200	4000	8000	32,000	nd	200	-
Girard Point	200	80	2000	8000	8000	nd	200	-
Prospect Hill	-	-	-	20	-	nd	4000	-
Puumala (NE)	-	-	-	-	-	nd	-	1280

Titers expressed as reciprocal of the highest dilution of antisera with greater than 80% reduction of approximately 100 plaques.
- Titer less than 20
nd = not done

SCHMALJOHN, HASTY & DALRYMPLE

To establish conclusively that Hantaviruses conformed to the most stringent criteria of Bunyaviridae, that is, that all viruses possessed a conserved 3´terminal nucleotide sequence which was the same as the prototype virus Hantaan, we performed sequence analyses of each of the RNA segments of the eight isolates. These sequence results, which are summarized in table 5, indicated that all viruses within the proposed Hantavirus genus which we examined did possess a conserved sequence on each RNA segment. As expected, eventual sequence divergence among virus isolates as well as between segments of RNA of individual viruses was apparent and began at the 14th nucleotide from the 3´ terminus. These data, provide difinitive molecular evidence that the Hantaviruses are a unique group of Bunyaviridae.

Table 5.
3´ NUCLEOTIDE SEQUENCES OF HANTAVIRUS RNA SEGMENTS

VIRUS

LRNA

HANTAAN 3´	AUCAUCAUCUGAGGGAUUUAUUGAU	
LEE	••••••••••••CU•CAUG•UUG	Note: • indicates
URBAN RAT	••••••••••••CC••CUC••U•	nucleotide homology
TCHOUPITOULAS	••••••••••••CU•C•C•GUUA	with the prototype
GIRARD POINT	••••••••••••CU•CAUG•UGG	virus: Hantaan.
PROSPECT HILL	•••••••••••XCUC•AUXXXXX	
PUUMALA (NE)	••••••••••••CU•C•C•GUU•	
SAPPORO RAT	••••••••••••CC••CUC•CG•	

M RNA

HANTAAN 3´	AUCAUCAUCUGAGGCGUUUUCUUUG
LEE	•••••••••••••••••••••A•••C
URBAN RAT	•••••••••••••••••C•U•G•C
TCHOUPITOULAS	•••••••••••••••••C•U•G•C
GIRARD POINT	••••••••••••••••••••••••C
PROSPECT HILL	••••••••••••••••A••CU••CX
PUUMALA (NE)	•••••••••••••••••C•U•G•C
SAPPORO RAT	•••••••••••••••••C•U•G•C

S RNA

HANTAAN 3´	AUCAUCAUCUGAGGGAUUUCUCGAU
LEE	•••••••••••••••••••••••••
URBAN RAT	•••••••••••••••••••••••••
TCHOUPITOULAS	•••••••••••••••••••••••••
GIRARD POINT	•••••••••••••••••••••U••X
PROSPECT HILL	••••••••••A•C•••••••••••
PUUMALA (NE)	•••••••••••••••••••••••••
SAPPORO RAT	•••••••••••••••••••••••••

In summary we have examined eight virus isolates some of which must be considered a potential health threat to troops stationed throughout the Eurasian continent. The viruses examined represented widely diverse geographical areas of isolation as well as the known host range of these agents. Among these viruses are several which are suspected or are known etiological agents of at least three forms of hemorrhagic fever with renal syndrome. Our data have indicated that all viruses are antigenically and molecularly related but are not identical to each other or related to any other known group of viruses. While we feel that these viruses should be classified as a fifth genus of the Bunyaviridae, they are unique from other Bunyaviridae not only molecularly and antigenically, but also biologically. Unlike typical Bunyaviridae, no arthropod vectors have yet been detected for the Hantaviruses. Transmission to humans is believed to occur directly from rodents, via aerosolization of saliva, urine and feces. The large number of cases of Korean hemorrhagic fever which occurred during the Korean conflict was most likely related to this transmission route when troops involved in digging trenches unearthed Apodemus burrows and consequently created aerosols.

Also in contrast to the Bunyaviridae, these viruses have the extreme propensity to persist in their rodent hosts long after infection. It is believed that once infected, an Apodemus may continue to excrete infectious virus in its urine for life. Recent studies have indicated that viruses related to Hantaan virus are persistent in rodents throughout the world (Maj. J. LeDuc personal communication). Why these viruses sometimes emerge from a harmless rodent virus to a deadly human infectious agent is not known. It is apparently not simply a host-range phenomenon because, as we have shown, there is a marked antigenic cross-reactivity of the viruses from similar hosts and human disease has been associated with viruses from all rodent hosts.

One possibile explanation is that only certain genes or combinations of genes from different viruses might result in a virus more virulent for humans. As we have demonstrated, the genomes of Hantaviruses isolated in the same endemic disease area (e.g. Hantaan and Lee) as well as those isolated from non-endemic disease areas (e.g. Tchoupitoulas and Girard Point) are molecularly quite diverse. Other Bunyaviridae have been reported to be able to exchange genome segments with serologically related viruses, and consequently select for the most virulent gene combination. We have not yet been able to examine this possibility with Hantaviruses, and await production of recombinant DNA clones which will be useful not only for studying the replication of these viruses but should also provide precise probes for examining important gene regions. Of greatest interest will be genes which might be responsible for virulence and also those regions which code for antigens involved in viral neutralization and induction of protective antibodies.

One of the most important questions for those of us involved in trying to prevent as well as understand virus infections and disease, is whether or not to investigate live attenuated vaccines against agents such as Hantaan virus, which shows such a pronounced persistence in vivo. The long-range effects of persistent viral infections must be of concern. Perhaps the best approach to a vaccine is production of protective synthetic peptides or recombinant DNA products. Alternatively, the feasibility of using antigenically related rodent viruses with no known human disease potential as a vaccine might be explored.

While we do not yet know the answers to these questions, our findings constitute a significant first step in the understanding of these elusive agents, the diagnosis of infection and the development of methods for preventing infection and disease.

REFERENCES

1. WHO REGIONAL OFFICE FOR THE WESTERN PACIFIC (1982). Report of the working group on haemorrhagic fever with renal syndrome. Tokyo. WPR/RPD/WG/82.16.

2. LEE, H.W., LEE, P.W., and JOHNSON, K.M. (1978). Isolation of the etiologic agent of Korean hemorrhagic fever. J. Infect. Dis. 137, 298-308.

3. FRENCH, G.R., FOULKE, R.S., BRAND, O.A., EDDY, G.A., LEE, H.W., LEE, P.W. (1981). Korean hemorrhagic fever: propagation of the etiologic agent in a cell live of human origin. Science 211:1046-1048.

4. SCHMALJOHN, C.S., HASTY, S.E., HARRISON, S.A., and DALRYMPLE, J.M. (1983). Characterization of Hantaan virions: Prototype virus of hemorrhagic fever with renal syndrome. J. Infect. Dis. 148, 1005-1012.

5. SCHMALJOHN, C.S., and DALRYMPLE, J.M. (1983). Analysis of Hantaan virus RNA: Evidence for a new genus of Bunyaviridae. Virology 131, 482-491.

6. SCHMALJOHN, C.S. and DALRYMPLE, J.M. (1983). Biochemical characterization of Hantaan virus. In "The Molecular Biology of the Negative Strand Viruses." Academic Press.

7. DE WACHTER, R., and FIERS, W. (1972). Preparative two-dimensional polyacrylamide gel electrophoresis of ^{32}P-labeled RNA. Anal. Biochem. 49, 184-197.

8. CLERX-VAN HAASTER, C.M., CLERX, J.P.M., USHIJIMI, H., AKASHI, H., FULLER, F. and BISHOP, D.H.L. (1982). The 3' terminal RNA sequences of Bunyaviruses and Nairoviruses (Bunyaviridae): Evidence of end sequence generic differences within the virus family. J. Gen. Virol. 61, 289-282.

9. PEATTIE, D.A. (1979). Direct chemical method for sequencing RNA. Proc. Nat. Acad. Sci. USA 76, 1760-1764.

10. ZOLLINGER, W.D., DALRYMPLE, J.M. and ARTENSTEIN, M.S. (1976). Analysis of parameters affecting the solid phase radioimmunoassay quantitation of antibody to meningococcol antigens. J. Immunol. 1976, 1788-1798.

RESOURCE ALLOCATION: MANNING THE FORCE IN THE 80'S

ROBERT C. SCHWALM
US ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES
FT BLISS, TX   79916

The ever-increasing high-tech manning requirements of the Army's new and emerging weapon systems require that the Army make optimal use of a shrinking manpower pool in support of those systems. It is no longer enough simply to throw more hands at a problem; it is now necessary that the right hands be applied. This is especially the case with the more sophisticated weapon systems, where the technology of the system requires particular personnel. However, even the less advanced systems must take into account the personnel assigned, if only because fewer personnel tend to be available these days.

The force structure issue has become one of assigning the right personnel in the proper numbers to guarantee efficient and effective operation of a weapon system. The problem is how best to determine the right numbers of the right personnel, especially when the system being studied is only slightly further along than concept development. Put another way, how does the Army structure its forces for maximum potential from minimum resources?

One solution--if the system already exists and money and personnel are no problem--would be to actually test force structure alternatives out in the field. Unfortunately, this solution is seldom feasible.

Because of this, while force structure issues have not been left to chance, they have not been addressed by the Army in any rigorous or systematic way. Most often, force structure decisions are based upon precedent, experience, and the opinions of subject-matter experts. Additionally, there are numerous considerations--some political and some pragmatic--which are not directly related to force structure issues but which most definitely impact on force structure decisions: These considerations include mental inertia--the basic urge to keep things the same; unit parochialism--the desire to maintain and protect one's resources; hedging-- planning and conserving resources to protect against a rainy day or future reductions in resources; and coping--planning around the fact that no

matter what the Army's intentions as stated in, say, a Table of Organization and Equipment (TOE), there are seldom as many people or equipment items as specified, and still fewer when attrition due to schooling, health profiles, and maintenance are considered.

DEVELOPMENT OF AN APPROACH

To address force structure issues in light of the current manpower shortfall, the Army must develop a more rigorous and systematic approach-- one which protects essential resources under all conditions and can be applied evenly and comprehensively across systems and across time. The approach, then, must be based upon the objective and the quantifiable and at the same time must take into account the relative priorities of certain skills, missions, functions, personnel, and equipment. In all likelihood no single instrument--whether a computer-based model or paper-and-pencil algorithm--will solve all force structure problems at all levels of the force. Instead the Army must move toward a combination of tools applied systematically and comprehensively to force structure issues.

At present there are a variety of decision aids available to help those in force design, but most of these aids--when used at all--are applied haphazardly and inconsistently. For the most part the primary decision tool in force design remains the advice of subject-matter experts. But the demands of projected force considerations make that tool at best a secondary resource. The Army is essentially being forced by events to turn to more objective and comprehensive information bases. A methodology based on computer-simulation models would appear to provide this base.

With this in mind, a much more effective solution--in terms of cost, time, personnel, and materiel--is to simulate the force structure alternatives with the use of computer models. The US Army Research Institute for the Behavioral and Social Sciences (ARI) has initiated a research effort in this area, applying computer-based models to the efficient assignment of tasks, personnel, and equipment to crew-served weapon systems and unit force structure.

MODELS ADDRESSING RESOURCE ALLOCATION

The effort began with the development of the ARI Crew Performance Model (CPM). Dr. Lloyd M. Crumley of the ARI Scientific Coordination Office at Fort Sill recognized that the Army had no easy means for determining the impact of crew factors (e.g., crew size or task assignment) on crew performance. Additionally, any efforts at determining the impact of such factors were likely to be based on subjective experiences rather than on empirical data. Dr. Crumley developed the Crew Performance Model to help correct these shortcomings and to address other performance considerations.(1)

In its simplest form, the CPM simulates the effects of changes in crew size and/or the assignment of tasks on crew performance.(2) Crew performance can be measured as the time to complete an activity (e.g., emplace a howtizer) or as the amount of involuntary downtime per crew member during the completion of the activity. Together the measures provide an estimate of the relative efficiency of a crew structure.

The CPM consists of three components: a) the task library, which lists the tasks to be performed, the times required to perform the tasks, and any constraints on crew members performing the tasks; b) the task structure, in which the user specifies which crew member is to perform which tasks and in what order; and c) the main program, which accesses the task information contained in the task library to simulate the task structure as outlined. Once the user has developed the task library, it becomes a simple matter to make changes in the task structure to determine the effects of those changes on crew performance.

The second part of ARI's research program on resource allocation involves the application of the Analysis of Military Organizational Effectiveness (AMORE) methodology to force structure issues. AMORE considers the effect on unit capability of degrading factors as a function of time. Unit capability is usually measured in terms of the capability of a unit to reconstitute itself after degradation. Degradation can take the form among others of casualties or losses after combat, the impact of wearing (even the threat of wearing) MOPP gear, or attrition due to drug abuse. AMORE can also be used with capability specified, allowing the user to vary resource requirements needed to achieve that level of capability. Input to the model includes initial unit resources (usually the TOE), substitution matrices for personnel and materiel, mission essential teams (the smallest increment or unit needed to complete the mission), and probabilities of degradation (the likelihood of damage to resources). From the model the user can determine not only the capability of the unit to recover from degradation, but also the skill positions and materiel which cause the model to "choke," i.e., shortfalls which prevent the model from building any further teams or reconstituting to higher levels. Finally, the model can be used to identify substitutions or areas for potential cross-training.

Although the CPM and AMORE can and have been applied independently, the natures and levels of their analyses make them eminently compatible. In a stair-step fashion, the results of the CPM analysis provide a context or framework for applying the AMORE methodology. Whereas the CPM considers the individual as part of a crew, AMORE considers the individual and crew as part of a larger unit, usually a battery. The CPM can provide insight into the relative strengths and weaknesses of various crew structures; AMORE considers the impact of those factors on the larger unit. In other

words, the CPM can be used to establish an optimal crew size; AMORE can then be used to build an optimal higher-level unit structure around that crew.

APPLICATION OF A FORCE STRUCTURE METHODOLOGY

By way of example, consider ARI's modeling effort on the M109 155mm SP howitzer system. Applying the stair-step approach described previously, ARI first performed a crew performance analysis on M109 howitzer crews. Using information and findings from that analysis as a springboard or framework, the AMORE methodology was then applied to an analysis of battery organization.

Specifically, the analysis proceeded as follows. In applying the Crew Performance Model, first a detailed task list was developed--defining the tasks, identifying constraints on performance of the tasks, specifying the times required to complete the tasks.(3) (The latter, by the way, were derived largely from time and motion analyses of crew operations.) Alternative crew structures--that is, variations in crew size and the assignment of tasks--were also developed; the specific alternatives were designed to promote efficient completion of the activities being modeled (in this case, emplacing, firing, and march-ordering the howitzer). Crews ranging in size from ten members (base case) down to four were compared.

With this information the model was exercised and the crew structures compared.(4) Not surprisingly, by reducing involuntary downtime through the more efficient allocation of tasks, completion time during all activities remains stable despite the decrease in personnel. Put another way, a reduction in personnel on the gun is compensated for by the increase in efficiency. This trend obtains through a reduced crew size of five people; at that point completion time begins to suffer.

It must be emphasized that these findings are not to suggest that a howitzer crew be reduced in size. Quite the contrary, our analyses suggest that not only are there enough support duties (e.g., ammunition resupply, sleep, perimeter defense) to keep a full ten-man crew busy, but that an 11th crew member is warranted. Still, at any given time only five or perhaps six men are needed to operate the weapon without a loss in completion time; the remainder, however, are necessary to perform required support duties off the gun.

It is interesting to note that many section chiefs already split their crews in just such a fashion. Because they seldom have the full complement of their ten-man crews, they have been forced to operate more efficiently both on and off the gun. The results of the crew performance analysis, then, provide some empirical justification for such split-crew procedures.

Given this information as background, the AMORE methodology was applied to the Division-86 M109 howitzer battery.(5) Although the analysis was intended primarily to examine the impact of crew size on battery capability (as had been the purpose of the previous CPM analysis), also considered were the individual and combined influences of four other factors--degradation level (casualties, from 10% to 40%), battery size (139 soldiers vs the proposed level of 129), number of howitzer sections per platoon (i.e., whether to break a battery into platoons at two or three sections), and, finally, MOSC changes (specifically, the introduction of 64Cs for certain 13Bs). (Because none of the latter three variables had a significant effect individually or in combination, only the effects of alternative battery (crew) organizations and degradation levels will be discussed.)

Six alternative howitzer crew structures were compared with the standard ten-man crew base case. The alternatives ranged from a mission-essential four-man crew (with six personnel in support roles) to a 14-man crew (the hypothetical crew size required for a howitzer crew to fully resource all its duties both on and off the weapon).

Basically it was determined that simply adding more hand⌐ not improve the capability of the howitzer battery to recover and .ᴄⱼᴄⱼte after casualties. Too often specific skills (e.g., gunner) w⌐ needed to rebuild teams. Thus, neither the 14-man crew nor the standard ten-m⌐n crew showed the capability, sustainability, or reconstitutabilit⌐ the split-unit crews. This became particularly evident as the casualty rate was increased: When larger crews are considered mission-essential, the impact of increasing casualties will be greater for those larger crews than for small crews. Thus, the 5/5 split crew (five men on the gun, five in support), particularly with an 11th crew member considered (5/5/1), was found to be especially robust.

The AMORE analysis, therefore, provided independent confirmation of the prior crew performance findings. It extended those findings, however, by considering the impact of possible crew changes, in addition to other factors, on the capability of the larger unit, in this case the battery.

CONCLUSION

Both the Crew Performance Model and AMORE have been successfully applied to a variety of military systems. Individually or together the models have been applied to such widely divergent systems as the M198 howitzer, Pershing II, Division Support Weapons System (DSWS), and Sgt York/DIVAD. The approach has been applied to units varying in size--from sections to platoons and batteries and divisions--and to units varying in stage of development (e.g., DSWS, M109). Perhaps most important, both models have been readily applied by users outside the developing agencies;

users at the level of the TRADOC proponent schools have had success in learning and operating the models and interpreting their output.

Although neither the CPM nor AMORE is a panacea for problems in force development/force structure, both are proven to have some power as empirical tools for addressing such issues. Furthermore, both models in particular and the approach in general are being improved to provide better analyses with enhanced realism and wider applicability. However, these are only two tools--really, two first steps--in the development of a competent resource allocation methodology.

REFERENCES

1. Crumley, L. M. Preliminary plan for a potential USAFAS sponsored study of the effect of long term, continuous duty on howitzer crew performance. ARI Ft Sill Field Unit Working Paper 79-2, 1979.

2. Schwalm, R. C., Crumley, L. M., Coke, J. S., & Sachs, S. A. A description of the ARI Crew Performance Model (RR 1324). Alexandria, VA: US Army Research Institute for the Behavioral and Social Sciences, 1981.

3. Coke, J. S., Crumley, L. M., & Schwalm, R. C. Emplacing, firing and march-ordering an M109A1 howitzer: Tasks and task times (RR 1312). Alexandria, VA: US Army Research Institute for the Behavioral and Social Sciences, 1981.

4. Crumley, L. M., Schwalm, R. C., & Coke, J. S. An evaluation of the effects of various task assignment alternatives on M109A1 howitzer crew performance (RR 1337). Alexandria, VA: US Army Research Institute for the Behavioral and Social Sciences, 1982.

5. Robinson, R. E., & Hannon, C. W. An analysis of the capability of alternative Division-86 howitzer battery organizations (TR). Alexandria, VA: US Army Research Institute for the Behavioral and Social Siences, in preparation.

SPIN-UP FROM REST IN A LIQUID PAYLOAD (U)

*RAYMOND SEDNEY, DR., and NATHAN GERBER, MR.
U.S. ARMY ARMAMENT RESEARCH AND DEVELOPMENT CENTER
BALLISTIC RESEARCH LABORATORY
ABERDEEN PROVING GROUND, MARYLAND 21005

I. INTRODUCTION

For many years it has been observed that liquid-filled projectiles have a proclivity for unusual flight behavior, often being unstable even though the same projectile with a solid payload is stable. The essential mechanisms for this behavior are rotation imparted to the liquid by the spin of the projectile and the ability of a rotating fluid to support waves, first called inertial waves by Kelvin. The problem has been studied extensively for the case of "solid body rotation," i.e., late enough in the flight so that the liquid is fully spun-up. It was found that projectile instability was caused by resonance between an inertial oscillation in the liquid and the coning motion of the projectile. However, projectile instabilities have been observed early in flight, while the liquid is in the transient state of motion called spin-up. Indeed, the parameters of the gun-projectile system may be such that the liquid might not achieve solid body rotation during flight time of the projectile. Several new problems and phenomena are introduced when one attempts to find the frequencies and damping of the inertial waves that exist during spin-up. The determination of the transient spin-up flow and then the wave frequencies and damping are the subjects of this paper. The essential aspects are presented here; detailed discussions are given in References 1, 2, and 3.

The wave frequencies and damping are obtained from the solution of the free oscillation problem in the rotating fluid; this requires solution of an eigenvalue, abbreviated e.v., problem of considerable complexity. One of the interesting complexities is the "critical layer." The term critical layer is used because the phenomenon is analogous to the critical layer that occurs in the Orr-Sommerfeld equation which governs the stability of a laminar shear layer. For the rotating fluid the physical interpretation of the critical layer is that the wave frequency is an integral multiple of the local angular frequency of the spin-up flow, both quantities being functions of time; mathematically, the governing system of differential equations has a turning point. One can also interpret the critical layer

191

as a resonance, quite distinct from the resonance discussed in the first paragraph.

For the case of solid body rotation and inviscid perturbation thereon, Stewartson[4] showed that resonance between the wave frequencies in the liquid and the nutational frequency of the projectile angular motion causes instabilities; Wedemeyer[5] derived 'scous corrections to the theory. The idea of resonance can be and has been used beyond the restrictions of the Stewartson theory; it would then be proper to call it Stewartson's principle.* Thus, a necessary condition for instability is that the nutational frequency of the projectile is equal to a wave frequency at some time during spin-up. To determine if instability occurs, it is necessary to calculate the moment exerted by the liquid by solving the forced oscillation problem. An approach to this problem was presented by Murphy[6] for large time, i.e., close to solid body rotation, using inviscid perturbations; the calculation fails when the critical layer exists because viscous perturbations, as used here, are needed to resolve the critical layer.

Two brief digressions follow to put liquid-filled projectile problems in historical and scientific perspectives. Apparently, liquid payloads were used at least as early as the 1920's, see Reference 7. The erratic behavior of these projectiles remained mysterious until Stewartson's work appeared in a classified report in 1953 and in the open literature in 1959.[4] In the late 1950's Karpov initiated experimental research in the area at the BRL. This showed the need for viscous corrections to Stewartson's theory. Wedemeyer[5] developed these and showed remarkable agreement between his theory and the experimental results. In another basic work[8] Wedemeyer derived a model for the unperturbed spin-up flow. Knowledge had matured enough so that an AMC handbook was written by Karpov and Wedemeyer[7] in 1969. Recent work at BRL has concentrated on the spin-up problem.

Scientifically, liquid-filled projectile problems require the theory of rotating fluids.[9] This theory has application in astrophysics, geophysics, and various technologies. Wave motion in solid body rotation, spin-up, and, to a lesser degree, spin-down have been studied for these applications. Spin-down flow has not been studied extensively because, for a large range of the parameters, the flow is unstable. The observation of tea leaves collecting at the center of a cup after stirring is a familiar example of a stable spin-down flow. Knowledge of spin-up flow alone is not very useful for the projectile problem. Perturbations on this flow, to give the wave system, are needed; this area had not been studied in previous work on rotating fluids.

---

*This basic work by Prof. Keith Stewartson is the only paper he wrote on this subject, but he retained his interest in the work on liquid-filled projectiles at BRL until his untimely death in 1983.

Typically, the liquid is contained in a right circular cylinder. It is assumed that spin is imparted impulsively to the cylinder; the angular acceleration of the projectile in an artillery gun tube is about 0.03 sec^{-2}, the acceleration time scale is 0.005 sec and the time in bore is 0.020 sec; see Reference 1 for justification of the impulsive start assumption. Notation used here and the cylindrical polar coordinate system are shown in Figure 1. The spin-up flow is determined by only two non-dimensional parameters:

$$\text{Reynolds number} = Re = \Omega a^2/\nu \quad \text{and} \quad \text{aspect ratio} = A = c/a$$

where Ω is the spin (rad/sec), a and c are the radius and half-height of the cylinder, and ν is the kinematic viscosity of the fluid. The range of these parameters is large. Experimental data such as velocity fields, pressures, or gyroscopic motion of the container exist for $1 < Re < 10^7$ and $0.5 < A < 5$.

The analyses presented here are restricted to larger Re by virtue of some asymptotic approximations. The lower bound on Re is not known, in general. From Reference 10 it appears to be $O(10^2)$. The e.v. analysis is also limited because the boundary layers on the cylinder endwalls, called Ekman layers, are not included. Finally, the angular motion of the projectile must be restricted to small angles because the theory is linear.

Whether or not spin-up effects are important in a projectile flight, rather than solid body rotation, can be estimated by comparing the projectile time of flight, $\bar{t}_{fl}$, with the characteristic time for spin-up determined by

$$\bar{t}_s = (2 \ c/a) \ Re^{1/2}/\Omega \qquad \text{(sec)}$$

if the Ekman layers are laminar. For $Re > 10^5$, approximately, the Ekman layers may be turbulent, in which case the characteristic spin-up time can be estimated from

$$\bar{t}_{st} = (28.6 \ c/a) \ Re^{1/5}/\Omega \qquad \text{(sec)}.$$

If $\bar{t}_s$ or $\bar{t}_{st} \ll t_{fl}$, spin-up effects can be neglected and solid body rotation can be assumed. If $\bar{t}_s$ or $\bar{t}_{st} \cong \bar{t}_{fl}/10$ or larger, spin-up effects probably have to be considered. For an artillery shell $\bar{t}_{fl} = 40$ sec is typical. Non-dimensional spin-up times, $t_s = \Omega \ \bar{t}_s$ and $t_{st} = \Omega \ \bar{t}_{st}$, are also used.

To put these estimates in perspective, consider two cases which will be used in presenting illustrative numerical results (courtesy of Dr. W. P. D'Amico, BRL):

	Re	c/a	Ω (rad/sec)	ϵ_s or ϵ_{st}
Case 1:	4,974	3.30	8937	0.052
Case 2:	1.99×10^6	5.20	754	3.510

For Case 1, the parameters are appropriate to a test of small caliber projectiles in a ballistic range and for Case 2 to an artillery projectile. At these spin-up times the projectiles would be 48m and 1,230 m from the gun for Cases 1 and 2, respectively; in both cases observations on projectile motion could be made at these distances.

This paper is organized as follows: the physics and mathematical description of the spin-up flow are given plus some illustrative solutions. Perturbations on this flow are prescribed to determine the inertial waves and the required e.v. problem is formulated. The numerical methods used in solving this problem are outlined and the existence of the critical layer is discussed. Results for the e.v. and critical layer are presented and comparison with experimental data is shown.

Partial validation of the theory has been provided by experiments[11] and numerical simulations.[10] At the present time there are no data from projectile firings that can be used to validate this theory.

II. THE SPIN-UP FLOW

Consider the axisymmetric, time dependent motion of a fluid which fills a right-circular cylinder, initially at rest. At $t = 0$ it is impulsively brought to a constant angular velocity, Ω. Lengths, velocities, pressure and time are made nondimensional by a, aΩ, $\rho \, \Omega^2 a^2$ and Ω^{-1}, respectively, where where ρ is the liquid density. In the inertial frame cylindrical coordinates r, θ, z are used and velocity components are U, V, W; see Figure 1. Dimensionless time is t and derivatives are indicated by
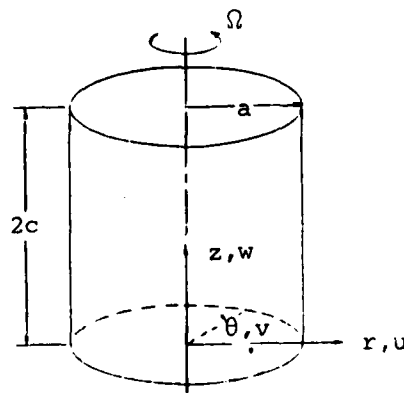
Figure 1. Notation for spinning Cylinder.

subscripts. The flow is governed by the Navier-Stokes equations; only the azimuthal, θ, momentum equation is shown here:

$$V_t^* + U^* V_r^* + W^* V_z^* + U^* V^*/r = Re^{-1} (\nabla^2 V^* - V^*/r^2) \qquad (2.1)$$

where the asterisk indicates the exact solution and ∇^2 is the Laplacian. The r and z momentum equations and continuity complete the set. In the recent past it became feasible to obtain finite difference solutions of these equations[12] for Re < 10^5. This was not the case when Wedemeyer[8] proposed his model for spin-up. Nevertheless, his model is still used as input to the e.v. problem because it would be impractical to use the finite difference solutions of the exact Navier-Stokes equations.

A. The Physics of Spin-Up. The validity of the Wedemeyer model can be discussed in terms of three time scales: the time for one revolution of the cylinder, $2\pi\Omega^{-1}$; t_s or t_{st}; and the time for vorticity (or velocity gradients) to diffuse radially, Re Ω^{-1}. The model requires $2\pi\Omega^{-1} \ll t_s \ll Re\ \Omega^{-1}$. It is known that the Ekman layers on the ends of the cylinder, Figure 2, form and become essentially steady in a time $2\pi\Omega^{-1}$. For small time and near the center of the endwalls these layers are essentially the same as the boundary layer on a steady, rotating disk, the von Karman disk problem; the rotating disk is a form of centrifugal pump. The suction exerted by these layers draws external fluid into them and imparts rotation to this fluid. With no pressure gradient acting, the fluid spirals out to larger radii where the Ekman layers eject it into the external flow. This fluid is now in the core, i.e., outside the Ekman layers. This is the basic mechanism for spin-up which takes place on a time scale t_s or t_{st}. Diffusion, on the time scale Re Ω^{-1}, takes longer by a factor of $O(Re^{1/2})$. In a plane θ = constant there is a recirculating flow as shown in Figure 2.
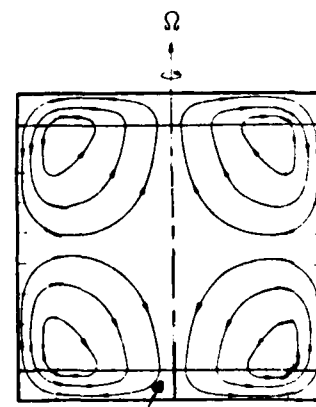
EKMAN LAYER

Figure 2. Schematic of Ekman Layers and Circulation in θ = Const. Plane During Spin-Up.

These qualitative aspects of the model are confirmed by the finite difference solutions to the Navier-Stokes equations. The quantitative agreement between predictions from the model and these solutions or experimental results is discussed in References 1 and 12.

B. Mathematical Description. Wedemeyer showed that the flow can be divided into two regions: the quasi-steady Ekman layers at the endwalls and the rest of the flow, called the core flow. He did not point out that a boundary layer, a Stewartson layer, must be inserted at the sidewall.[1] Order of magnitude arguments were used to simplify the Navier-Stokes equations in the core. For example, (2.1) is approximated by

$$V_t + U (V_r + V/r) = Re^{-1} [V_{rr} + (V/r)_r], \qquad (2.2)$$

and the results

$$U_z = V_z = P_z = 0 \qquad (2.3)$$

are also obtained; the asterisk is dropped to indicate this approximation. For large Re, Wedemeyer proposed neglecting the diffusion terms in (2.2), i.e, setting the right-hand side zero. This equation has a simple solution and was adequate for his purposes but it is not for the e.v. problem.[3]

To solve (2.2) a relationship between U and V is required. At this point Wedemeyer was forced to take a phenomenological approach. For laminar Ekman layers his result is

$$U = -\kappa (a/c) Re^{-1/2} (r - V) \qquad (2.4)$$

with $\kappa = 0.443$; Greenspan[9] suggested $\kappa = 0.5$. For turbulent Ekman layers his result is

$$U = -0.035 (a/c) Re^{-1/5} (r - V)^{8/5}. \qquad (2.5)$$

Note that the core flow is always assumed to be laminar so that turbulent stresses are not introduced in the right-hand side of (2.2). The relations (2.4) and (2.5) are called compatibility conditions because the Ekman layer suction must be made compatible with the core flow described by (2.2). Much more should be said about compatibility conditions but no further discussion can be given here; see Reference 1 for this and an explanation of the confusion that has appeared in the literature on this point.

Inserting (2.4) or (2.5) into (2.2), V can be determined from this nonlinear second order equation of diffusion type. It is integrated by finite difference methods using some standard techniques;[1] solving (2.2) is considerably easier than the Navier-Stokes equations. The boundary and initial conditions are $V = 0$ at $r = 0$, $V = 1$ at $r = 1$, and $V = 0$ at $t = 0$. Special treatment is required near the point $r = 1$, $t = 0$ because, for an impulsive start, a discontinuity in the boundary conditions exists there.[13] In most of our e.v. calculations we used the V from the numerical solution of (2.2) with either (2.4) or (2.5); the latter give U, and W is obtained from the continuity equation.

C. Results. Some examples of the solutions of (2.2) are given in Figure 3 for Cases 1 and 2 using (2.4) and (2.5), respectively. For each case, results for three non-dimensional times are shown. The nondimensional spin-up times are $t_s = 460$ and $t_{st} = 2646$ for Cases 1 and 2, respectively. As $t \to + 0$, V tends to a discontinuous function; for small t an asymptotic solution to (2.2) was derived in Reference 1. In the e.v. analysis a quasi-steady assumption is used which is violated as $t \to + 0$. A practical limit on how small t should be in the analysis is set by the smallest t at which data can be obtained. Comparisons of solutions of (2.2) and the Navier-Stokes equations[1] validate this model over a range of Re and A.
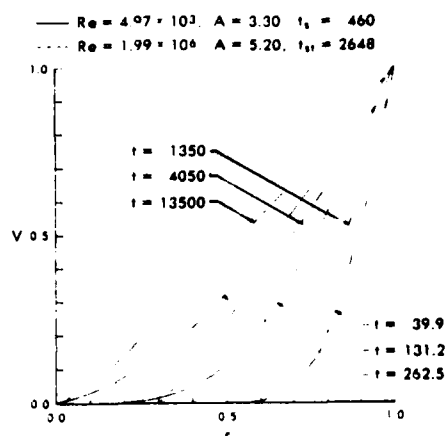
Figure 3. V vs r for Cases 1 and 2.

III. INERTIAL WAVES IN THE SPIN-UP FLOW

A. The Perturbed Flow. The procedure for deriving the equations governing the perturbed flow is a standard one and will only be outlined. The velocity components and pressure are expressed as the sum of the spin-up flow and perturbation, e.g., $\overset{*}{U} (r,z,t) + u' (r,\theta,z,t)$ and then substituted into the Navier-Stokes equations for 3-D, unsteady flow; the prime indicates perturbation. Note that the θ dependence of u' is essential in the perturbation problem. Initially, the spin-up flow is a solution to the

Navier-Stokes equations for axisymmetric flow, i.e., the zeroth order terms. The first order terms, linear in the perturbations, are retained and the second order terms neglected. The coefficients in the linear perturbation equations are the spin-up flow variables; these are now approximated by the results from the model, Section IIB. A significant simplification of the perturbation equations results: only V and V_r appear in the final equations. Furthermore, the variation of V with t is small compared to that for the perturbations and t can be regarded as a parameter in V (r,t). This is the quasi-steady approximation.

B. The Eigenvalue Problem. To obtain the required wave type solutions, the perturbations are represented as a superposition of modes or, equivalently a triple Fourier expansion in θ, $\tilde{z}$ and t with coefficients functions of r; here $\tilde{z} = z + A$. It is exceedingly convenient to use complex notation for the modal forms so that

$$u' = \text{Real} \left\{ u(r) \cos K\tilde{z} \exp \left[i \left(Ct - m\theta \right) \right] \right\} \qquad (3.1)$$

with similar expressions for v' and p'; w' has the same form except for a $\sin K\tilde{z}$ factor. Here $K = k\pi/2A$, so that $k = 1,2,\ldots,$ and $m = 0,\pm1,\ldots$ are the axial and azimuthal wave numbers, respectively; $m = 1$ is usually the value relevant to the projectile problem. The complex quantities u, v, w, p are the eigenfunctions; they are solutions of the ordinary differential equations obtained by substituting the modal forms, e.g., (3.1) , into the perturbation equations. The nondimensional complex constant $C = C_R + i C_I$ is the eigenvalue of the system. The dimensional wave frequency is $C_R\Omega$ and the damping is $C_I\Omega$.

For this free oscillation problem the boundary conditions at the endwalls $z' = 0$, 2A are the no-slip conditions $u' = v' = w' = 0$ if the complete flow is being perturbed. Since we are perturbing the core flow and not the Ekman layers, the boundary conditions are not the same. The modal forms give $w' = 0$, but u' and v' $\neq 0$. Neglecting the Ekman layer flow, i.e., not satisfying the no-slip conditions, causes an error of about 1% in C_R but a large error in C_I. This is to be expected since significant dissipation occurs in the Ekman layers. Of course, the primary interest in this work is the calculation of C_R which is not sensitive to the Ekman layer flow.

The differential equations for the eigenfunctions form a 6th order system in complex form (12th order in real form). The no-slip boundary conditions apply at the sidewall, $r = 1$. Many of the familiar properties of e.v. problems do not exist for this system because it is not self-adjoint, e.g., it is unclear how to define a radial mode in a general and

198

unambiguous manner. Nevertheless, an index n is used to order the spectrum on the assumption that it is not continuous, and the e.v. are denoted by C_n. Three other complications arise in the numerical solution of the e.v. problem. (i) The system of differential equations has a singularity at $r = 0$, the center of the cylinder. Analytic solutions must be used there. (ii) The coefficient of the highest order derivatives in the differential equations, Re^{-1}, is small for the values of Re of interest. This classifies them as stiff equations. A technique called orthonormalization is used to ensure that the solutions remain linearly independent as the integration proceeds from $r = 0$ to $r = 1$. (iii) When the critical layer exists, high frequency oscillations occur in the eigenfunctions; the integration scheme and the number of significant figures in the computation must be capable of resolving these. Other important details in the numerical method of solving the e.v. problem are discussed in References 2, 3, and 13.

A brief description of the critical layer theory[3] will be given here. Its definition is best appreciated by considering the inviscid limit of the perturbation equations, i.e., the limit $Re \to \infty$. The order of the system is reduced from six to two and the coefficient of the highest (second) order derivative is $C_n - mV/r$. If this quantity vanishes the inviscid perturbation equations have a singularity. To obtain a real root, consider $C_{Rn} - mV/r = 0$. If this equation has a real root, r_c, where $0 < r_c < 1$, r_c is called the critical level. The neighborhood of r_c is called the critical layer. The physical interpretation of $C_{Rn} = mV/r$ at $r = r_c$ is that the wave frequency, C_{Rn} is an integral multiple of the angular frequency, V/r, of the spin-up flow. which indicates resonance between the n-th radial mode of the inertial waves (for given k) and the circular motion of the spin-up flow. A critical layer always exists unless $m = 0$ or if m and C_R have opposite signs. There can be a critical layer for each n and k.

Returning to the viscous perturbations, r_c is not a singularity of them. Rather r_c is a turning point indicating that the nature of the solution for the eigenfunctions changes in the critical layer. For example, high frequency oscillations of large relative amplitude develop and rapid changes in phase of u, v, w, and p occur in the critical layer.

C. Results. In presenting some representative results it is convenient to designate the three wave numbers in the (z, r, θ) directions by the triplet (k, n, m). In Figure 4 the eigenfunction $w_R = \text{Real}(w)$ vs r is shown for three times: at $t = 7,000$ there is no critical layer; at $t =$

1,000 and r_c = 0.44, as indicated by the arrow in Figure 4a, a rapid variation of w_R exists; at t = 400 and r_c = 0.66 the high frequency oscillations are shown. It is remarkable that w_R has 20 zeros in Figure 4c.

Time histories for C_1 and C_2 are shown in Figure 5 for Case 1 and modes (3,1,1) and (3,2,1). There is a maximum in C_{R1} at t = 45. This is typical for all C_{R1} vs t curves. A maximum in the C_{R1} vs t curve is signficant because of the necessary condition for projectile instability discussed in the Introduction. If the nutational frequency is less than the maximum of C_{R1}, there are two times at which instability might develop; the implications of this result have not been considered in projectile firings. Additional comments on the results shown in Figure 5 are given in Reference 3.

Figure 5. C_R and C_I vs t for Case 1; n = 1,2.

Figure 4. Eigenfunction Real (w) vs r; Re = 5 x 10^5, A= 2.7, Mode (5,1,1).

The time history for C_{R1} is shown in Figure 6 for Case 2, mode (5,1,1). The maximum occurs for t < 1200 and is not shown. Calculations for this case and some related ones were used to plan projectile firing tests and then to analyze the results.[14] Sensitivity of C_{R1} to variations in the parameters must be considered. The results for two such variations are also shown in Figure 6; Re was decreased by a factor of 10 and A was increased by 10%. There is a large change in r_c (t) for the first variation but a small change for the second.

200

In Figure 7 r_c (t) is presented for Cases 1 and 2 and n =1 and 2. Note the two time scales. For both cases r_c = 0 at an earlier time for n = 1 than n = 2 and the critical layer exists for a substantial part of the projectile flight time. The results in Figure 7 show that the critical layers for n = 1 and 2 can exist simultaneously. Of course, there is a critical layer for each n.
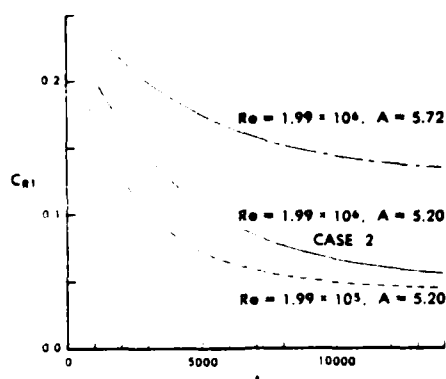
Figure 6. C_{R1} vs t for Case 2 and Two Variations of It.

Figure 7. r_c vs t for Cases 1 and 2; n = 1,2.

For the Orr-Sommerfeld equation a rapid change in phase of the velocity across the critical layer is known to exist, and it is crucial to the explanation of the insta- bility of a laminar shear layer. Even though the flow is stable in the rotating fluid case considered here, the flow variables u, v, w and p exhibit rapid phase changes across the critical layer. Results for β, the phase change in u, are shown in Figure 8 at three times and the value of r_c is indicated for each time. Another rapid change in β takes place in the sidewall boundary layer at r = 1. For t $\geq$ 10,000, r_c = 0 (see Figure 7) and β = constant except in the boundary layer. The phase change in p across

Figure 8. The Phase Angle of u' for Case 2 at Three Times.

the critical layer can provide the basis for an experimental determination of the critical layer effect; such an experiment is planned.

IV. DISCUSSION

In this paper the theory and a method for the solution of the spin-up e.v. problem were presented so that the inertial waves could be determined. Knowledge of these waves is required to study the stability of liquid-filled projectiles. The theory and method are successful in the sense that they provide results that are physically meaningful and do not violate intuition or the "physics of the problem." Other investigators have worked on this problem without success in that sense. Validation by comparison with numerical simulations, experimental results and projectile firings is required.

Limited validation using numerical simulation is provided in Reference 10. To the authors' knowledge the only reported measurements of C for m = 1 are in Reference 11. Some of these data are shown in Figure 9 for Re 44,412, A = 0.6 and mode (1,2,1). The calculated C_R is within the scatter of the data. This comparison validates the calculation of C_R for these parameters over the range 0.724 $\leq$ t/t$_s$ $\leq$ 2.252. For large t, essentially at solid body rotation, the experimental and calculated results differ by 1.5%.[11] The $r_c(t)$ curve shows that the critical layer exists over a considerable range for which data is presented. Comparisons for modes (2,1,0) and (1,1,1) give the same conclusion as for the (1,2,1) mode.
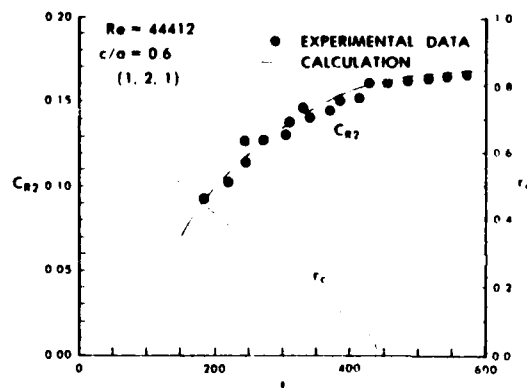
Figure 9. C_R vs t: Experimental Points and Calculation Values; r_c vs t.

t$_s$ = 253

There are no data from projectile firings that can be used to validate the theory presented here. Some indirect, qualitative results, obtained from yawsonde measurements, are given in Reference 14 which show consistency with the theory. Since no ballistic range tests are planned, further validation will probably depend on laboratory experiments.

REFERENCES

1. Sedney, R., and Gerber, N., "Viscous Effects in the Wedemeyer Model of Spin-Up From Rest," BRL Technical Report ARBRL-TR-02493, June 1983. (AD A129506)

2. Sedney, R., and Gerber, N., "Oscillations of a Liquid in a Rotating Cylinder: Part II. Spin-Up," BRL Technical Report ARBRL-TR-02489, May 1983. (AD A129094)

3. Sedney, R., and Gerber, N., "Numerical Study of the Critical Layer in a Rotating Fluid," AIAA Paper 84-0342, January 1984.

4. Stewartson, K., "On the Stability of a Spinning Top Containing Liquid," Journal of Fluid Mechanics, Vol. 5, Part 4, September 1959, pp. 577-592.

5. Wedemeyer, E. H., "Viscous Corrections to Stewartson's Stability Criterion," BRL Report No. 1325, June 1966. (AD 489687)

6. Murphy, C. H., "Moment Induced by Liquid Payload During Spin-Up Without a Critical Layer," AIAA Paper 84-0229, January 1984.

7. Engineering Design Handbook, Liquid-Filled Projectile Design, AMC Pamphlet 706-165, April 1969. (AD 853719)

8. Wedemeyer, E. H., "The Unsteady Flow Within a Spinning Cylinder," Journal of Fluid Mechanics, Vol. 20, Part 3, 1964, pp. 383-399.

9. Greenspan, H. P., The Theory of Rotating Fluids, Cambridge University Press, London and New York, 1968.

10. Sedney, R., Gerber, N., and Bartos, J. M., "Eigenfrequencies of Inertial Oscillations in a Rotating Fluid Via a Numerical Simulation," BRL Technical Report ARBRL-TR-02488, May 1983. (AD A129088). Also AIAA Paper 82-0296, January 1982.

11. Stergiopoulos, S., "An Experimental Study of Inertial Waves in a Fluid Contained in a Rotating Cylindrical Cavity During Spin-Up from Rest," Ph.D. Thesis, York University, Toronto, Ontario, February 1982.

12. Kitchens, C. W., Jr., "Navier-Stokes Solutions for Spin-Up from Rest in a Cylindrical Container," BRL Technical Report ARBRL-TR-02193, September 1979. (AD A077115). Also AIAA Journal, Vol. 18, No. 8, August 1980, pp. 929-934.

13. Sedney, R., and Gerber, N., "Treatment of the Discontinuity in the Spin-Up Problem with Impulsive Start," BRL Technical Report ARBRL-TR-02520, September 1983. (AD A133682)

14. D'Amico, W. P., Jr., "Flight Data on Liquid-Filled Shell for Spin-Up Instabilities," BRL Report in preparation. Also AIAA Paper 83-2143, August 1983.

CERAMIC ANODES FOR CATHODIC PROTECTION (U)

*ELLEN G. SEGAN, MS.
ASHOK KUMAR, DR.
MICHAEL OLSON, MR.
U.S. ARMY CONSTRUCTION ENGINEERING RESEARCH LABORATORY
CHAMPAIGN, IL 61820

Background

The corrosion of metallic structures immersed in water or buried in soil can be stopped by cathodic protection, i.e., by applying a small electric current from an outside source to the corroding structure. Cathodic protection has been used since 1824 when Sir Humphrey Davy introduced it as a means of protecting copper sheathing on ships.

Traditionally, impressed current cathodic protection systems have used anodes that are either inexpensive and very large, or small and expensive. High-silicon, chromium-bearing cast iron (HSCBCI) and graphite anodes often weigh more than 60 lb (24 kg). These anodes have posed many installation and maintenance problems, and are often damaged by ice and debris on hydraulic structures. An alternative system to the heavy anodes is the use of platinized anodes. These consist of a thin platinum layer on a passive substrate like niobium or titanium. They are much lighter because the anodic dissolution rate of platinum is very low; for example, the dissolution rate of platinum is less than one ten-thousandth that of HSCBCI. However, platinized anodes are vulnerable to abrasion and erosion corrosion damage, and the platinum makes them very expensive. Because of these problems, there is an acute need for improved cathodic protection systems with affordable anodes that can provide reliable protection with a minimum of installation and maintenance problems.

The possibility of using electronically conducting ceramics for anodes has been investigated (1,2,3). Many conducting oxides have been successfully used as electrode materials in high-temperature electrochemical cells and anodes for electrochemical processing. However, the use of ceramics for cathodic protection anodes has only recently been examined. Some oxide ceramics, such as ferrites, titanates, and cobaltates as well as some carbides and nitrides, exhibit semiconducting

behavior at ambient temperatures. Often, the conductivities of these ceramics approach that of impure metals. The semiconducting ceramics are largely made of inexpensive and abundant raw materials. Anodes coated with these ceramics contain a large proportion of oxygen, carbon, or nitrogen and a relatively low proportion of expensive or strategic metals. The expense of making ceramics at the high temperatures usually required has been greatly reduced by the development of processing techniques which minimize energy expenditures. These factors make ceramic anodes a promising option for impressed current cathodic protection systems.

Objective

The objectives of this investigation were: (a) to evaluate the feasibility of using electronically conducting ceramics for anodes in impressed current cathodic protection systems and identify applications where ceramic anodes can offer improved performance, and (b) to manufacture and test sintered ceramic anodes and anodes coated with low-resistivity ceramics that have low anodic dissolution rates and provide characteristics such as freedom of anode configuration, easy installation and replacement, small size, and toughness.

Electrical and Electrochemical Properties of Conducting Ceramics

Many ceramic materials exhibit high electrical conductivities because of their semiconducting properties. Semiconductors are generally classified as materials in which the electrical conductivity is electronic in nature and has values in the range of 10^{-3} to 10^{-9} ohm-cm. This is between that of metallic conductors (about 10^{-6} ohm-cm) and insulators ($>10^{12}$ ohm-cm).

The total electrical conductivity of any material is the sum of the conductivities of its individual charge carriers. In most ceramics, electrical conductivity is largely due to ionic conductivity, in which anions, such as O^{-2}, pass through the structure (crystalline or amorphous) by diffusion in an applied electric field. The mobility of cations and anions in ceramics is very low at ambient temperatures and increases significantly only at high temperatures (usually above one half of the melting point in degrees Kelvin). Most ceramic materials are covalently bonded, which causes localization of electrons in the structure. This results in very low electronic conductivity because electrons are "pinned" to their sites in the ceramic. At ambient temperatures, the ionic and electronic conductivities of most ceramics are very low. Thus, the traditional use of ceramics has been for insulators.

Some ceramic materials exhibit semiconducting behavior. This occurs because in some crystal structures, electrons may become mobile by transferring from one cation site to the next. This is a complicated process that gives them other unique magnetic and optical properties. Semiconduction is common in oxides that crystallize in spinel, hematite, and perovskite structures as well as in some carbides and nitrides. At ambient temperatures, ionic conduction in these materials is usually low in comparison to electronic conductivity.

Conducting ceramics have been used for some time as electrodes in high-temperature electrochemical cells and in chemical processing reactors. However, the use of ceramics for impressed current cathodic protection system anodes was only begun recently. Two types of systems have been tested. The first is a sintered ceramic anode, either with or without a metal core; the second is a plasma-sprayed ceramic on substrates such as titanium, niobium, or tantalum. Table 1 gives the dissolution rates of different anode materials.

Table 1

Typical Properties of Impressed Current
and Sacrificial Anodes

Electrode Material	Sacrificial or Impressed Current Anode	Anodic Dissolution Rate g/ampere year	Environment	Current Density A/m^2
Aluminum	Sacrificial	3,200	Soil and Water	
Zinc	Sacrificial	11,200	Soil and Water	
Magnesium	Sacrificial	7,900	Soil and Water	
Scrap Steel	Impressed	15-20	Soil and Water	
Scrap Cast Iron	Impressed	10-15	Soil and Water	
Silicon Iron	Impressed	1-2	Soil and Water	
Graphite	Impressed	2	Soil and Water	
Graphite	Impressed	200	3% NaCl Solution	500
Cast Magnetite	Impressed	40-50	3% NaCl Solution	500
Platinum-Coated Titanium	Impressed	0.01	3% NaCl Solution	500
Nickel Ferrite	Impressed	0.40-1.56	3% NaCl Solution	500
Magnesium Ferrite	Impressed	3.47	3% NaCl Solution	500
Zinc Ferrite	Impressed	3.28	3% NaCl Solution	500
Manganese Ferrite	Impressed	2.67	3% NaCl Solution	500
Cobalt Ferrite	Impressed	2.19	3% NaCl Solution	500

Magnetite anodes having a dissolution rate of 40 g per ampere year have been used for impressed current cathodic protection systems (1). These anodes exhibit chemical inhomogeneity, high porosity, and low mechanical strength. The density and mechanical strength can be improved somewhat by sintering NiO with Fe_2O_3, forming a spinel nickel ferrite. Best results have been achieved by using 40 mol percent NiO or 60 mol percent Fe_2O_3, thereby reducing the dissolution rate to 0.4 g per ampere year. However, these anodes do not have the required toughness and do not yield freedom of configuration for environments like locks and miter gates.

Typical properties of ferrite and cast magnetite sintered anodes investigated by Wakabayashi and Aoki (1) are summarized in Table 1. Nominal ferrite compositions were 0.1 MO - 0.9 Fe_2O_3, where M represents a divalent ion such as magnesium, zinc, manganese, cobalt, nickel, or iron. The table shows that nickel ferrite offered the lowest dissolution rate (1.56 g per ampere year). Wakabayashi and Aoki also found that the dissolution rate was reduced to 0.4 g per ampere year when nickel ferrous ferrite containing 60 mol percent Fe_2O_3 and 40 mol percent NiO was used. The high resistivity (0.3 ohm-cm) of nickel ferrite containing 40 percent NiO is acceptable in cathodic protection, because the anode resistance can be reduced by using special geometrical modifications such as tubular anodes.

Sintered ceramic anodes have good corrosion resistance and low resistivities. However, the toughness of the sintered ferrite anodes is low and is not acceptable in many applications where the anode can be mechanically damaged. The anodic dissolution rates of these anodes increase with decreasing current density and with decreasing chloride content in the water. The dissolution rate is also relatively constant in the pH range of 1 to 10, increasing markedly at pH < 1.

The marginal properties of sintered ceramic anodes have led to the development of ceramic-coated anodes. Various attempts have been made to manufacture anodes coated with magnetite. Itai and Kanai (2) used electrodeposition to coat titanium and other substrates with magnetite. The coating techniques employed electrodeposition of iron, dipping the iron-deposited titanium substrate into a solution of ammonium ferric oxalate under a reduced pressure, and then heating the treated substrate in an atmosphere of hydrogen and steam. The electrodes are suitable for producing chlorine and chlorates and for electrowinning metals. However, the maximum coating thickness obtained was only 20 microns; this does not meet the long-life requirement (10 to 20 years) of cathodic protection systems.

In another approach, Fujii, et al. (3), plasma-sprayed several 50-micron (0.002-in.) coatings of several spinel ferrites (magnetite, nickel, and cobalt ferrites) on titanium substrates and determined their anodic behaviors in sodium chloride solutions. The dissolution rates of the ferrite-coated anodes ranged from 0.1 to 8.7 g per ampere year, which is comparable to that of sintered ferrites. However, adhesion of the ferrite films to the substrate metals was not satisfactory. The process used reducing atmospheres during plasma-spraying, and the coatings produced were rather thin (50 microns [0.002 in.]). Plasma-spraying with 10 percent titanium dioxide mixtures or using a tantalum undercoating did not improve adhesion of the coating, and a workable anode could not be produced.

Optimization of Ceramic Anode Cathodic Protection Systems

Ceramic anodes show promise for use in impressed current cathodic protection systems. The sintered anodes appear to be usable in systems where the possibility of mechanical damage is small (i.e., buried pipelines and structures), while ceramic-coated anodes with high toughness can be used in environments where mechanical damage can occur (i.e., lock gates and the interior of water tanks). Good ceramic coatings can be prepared by a variety of techniques. Recent developments in plasma-spraying, ion plating, chemical vapor deposition, hot pressing, and other techniques can produce thick, adhesive coatings.

Laboratory Tests

Ceramic anodes were manufactured for the U.S. Army Construction Engineering Research Laboratory by plasma-spraying lithium ferrite on titanium and niobium substrates. Lithium ferrite raw materials were produced by grinding hot-pressed bars; the materials were deposited by plasma-spraying at a power level at 26.25 kW, with an argon forming gas at a flow rate of 1.7 m^3 per hour. Spray distance was maintained at 8.89 cm. The powder gas used was oxygen at a flow rate of 0.226 m^3 per hour. Strong, fine-grained adherent coatings of various thicknesses were deposited on titanium and niobium substrates. The substrates were machined out of solid rods.

Figure 1 shows the anode configuration used for testing the anodes in the field. The ceramic-coated anode consists of a plasma-sprayed lithium ferrite coating having an active surface area of about 0.001 m^2, 500 to 2500 microns thick, deposited on niobium or titanium substrates that are machined to a button shape, 2.54 cm thick, and threaded to fit a plastic gland. The ceramic coating faces the water side of the structure and the cathodic protection cable is connected to the metallic plug which is threaded into the substrate. The ceramic-coated anode is installed by drilling a hole in the steel structure to be protected and attaching the anode assembly.

The lithium ferrite coatings were chemically analyzed by Chicago Spectro Service Laboratory. The composition of the plasma-sprayed materials was determined by x-ray fluorescence of a fused sample. Lithium content was determined by atomic absorption because of its low molecular weight. The dissolution rates of the anodes were determined in the laboratory by immersing the anodes for 1 to 2 weeks at various anodic current loadings, both in tap water and in distilled water containing 3.5 percent reagent grade sodium chloride.

The amount of dissolved lithium and iron in the solutions was determined by atomic absorption analysis. Long-range tests of lithium

WEEP HOLE

PLASMA SPRAYED LITHIUM FERRITE COATING

TITANIUM OR NIOBIUM SUBSTRATE

METALLIC THREADED PLUG

PLASTIC GLAND

SEALANT

CATHODICALLY PROTECTED STRUCTURE

WASHER

NUT

NYLON FERRULE AND PLUG

CATHODIC PROTECTION CABLE

ANODE LEAD
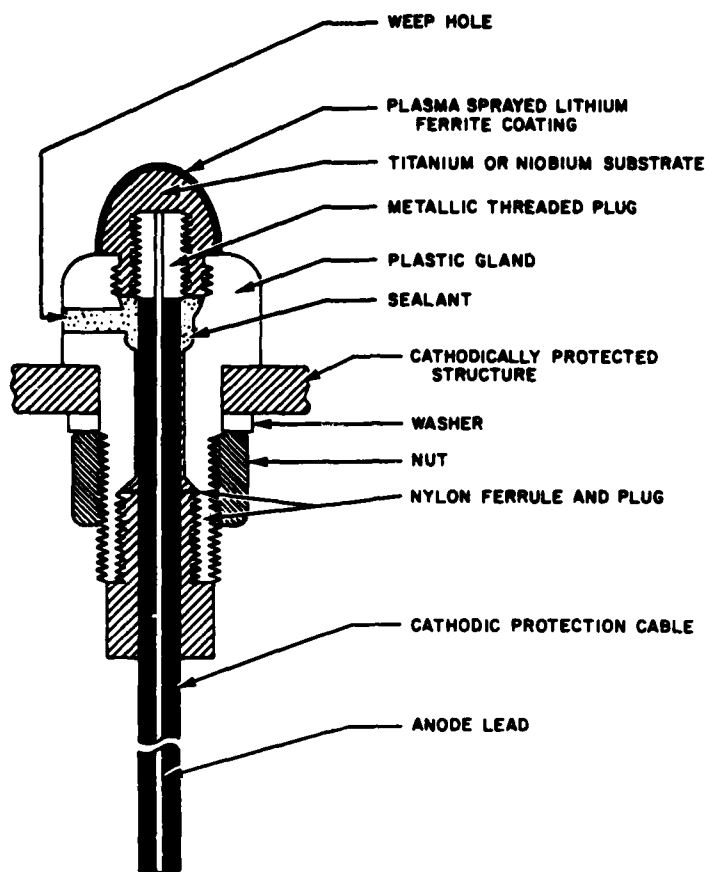
Figure 1. Ceramic-coated anode assembly.

ferrite were conducted by measuring the weight loss of an anode immersed in 3.5 percent sodium chloride in distilled water for 2 months at a current loading of 2000 A/m^2. The solution was replenished weekly.

In addition to plasma-sprayed lithium ferrite, several other sintered ceramics were produced and tested in a similar manner. These included selected donor doped and reduced alkaline earth titanates, NbO doped TiO, manganese ferrite, and manganese zinc ferrite. These materials were then subjected to a pass-fail anodic polarization test at various current levels in 3.5 percent reagent grade sodium chloride solution. Dissolution rates of materials which passed the initial test were determined by measuring the cation concentrations in solution by atomic absorption analysis.

210

Field Tests

Test anodes (25.4 mm in diameter) fabricated as shown in Figure 1 were installed at the Racine Lock gate structure on the Ohio River in West Virginia and at the Miller's Ferry Lock gate on the Alabama River in Alabama. Test anodes were also installed inside water storage tanks at Fort Eustis, Virginia, and at Fort Hunter Liggett, California. Anode assemblies were embedded in coke breeze contained in steel casings and installed to protect underground pipes at Fort Carson, Colorado, and at Fort Polk, Louisiana. The test anodes were installed, both in parallel and side by side, in existing cathodic protection systems designed with HSCBCI anodes. The current passing through the ceramic-coated anode and the applied voltage was measured.

Laboratory Tests on Ceramic-Coated Anodes

For coating thicknesses up to 30 mils, the plasma-sprayed lithium ferrite adhered to the niobium and titanium substrates for the duration of the tests. Thicker coatings were found to have extensive microcracking that apparently resulted from thermal expansion mismatches between the coating and substrate, which caused thermal cracking during cooling. Thus, it appears that the present coating process will have to be improved to produce thicker, nonporous coatings. Although anodes may have either niobium or titanium substrates, for this investigation, anodes made of 20 mils of lithium ferrite plasma-sprayed on titanium substrates were used unless otherwise noted.

Table 2 shows the composition of the plasma-sprayed lithium ferrite by both weight and mole percent. The oxygen content was estimated by assuming that the nonmetallic portion of the ceramic is oxygen. Analysis showed that the lithium ferrite contains significant amounts (molar proportions) of manganese, zinc, and calcium. These elements can all participate in the spinel structure on cation sites.

The chemical formula of stoichiometric lithium ferrite is $Li_{0.5}Fe_{2.5}O_4$, yielding an iron-to-lithium ratio of 5. The plasma-sprayed lithium ferrite had an iron-to-lithium ratio of 5.89, making it rich in iron. This is probably the result of preferential volatilization of lithium during the plasma-spraying process, since lithium is much less stable than iron at elevated temperatures. The chemical formula of the plasma-sprayed lithium ferrite can be determined by normalizing the proportions of the elements to yield four moles of oxygen. The following are formulas for some simple stoichiometric ferrites, and for the plasma-sprayed lithium ferrite:

Ferrite	$Fe^{3+}(Fe^{2+}Fe^{3+})O_4 = Fe_3O_4$
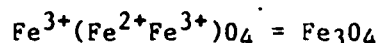
Spinel with + 2 substitution where $M^{2+} = Ca^{2+}$, Co^2, etc.
$Fe^{3+}(Fe^{3+}M^{2+})O_4$

Lithium ferrite
$Li_{0.5}Fe_{2.5}O_4$
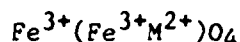
Plasma-sprayed lithium ferrite
$Li_{0.55}Fe_{3.25}O_4$

Plasma-sprayed lithium ferrite (major impurities included)
$Li_{0.55}Zn_{0.6}Mn_{0.14}Ca_{0.02}Fe_{3.25}O_4$
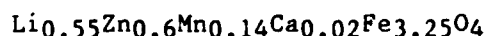
The plasma-sprayed lithium ferrite appeared to be oxygen-deficient, as shown by the high cation-to-anion ratio $(Li + Fe)/O_4 = 0.95$, compared to the stoichiometric ratio of 0.75. This effect is probably due to the reducing environment of the argon forming gas used during plasma spraying. Furthermore, the plasma-sprayed powders are deficient in lithium (note the low lithium-to-iron ratio -- 0.17 as compared to the stoichiometric ratio of 0.20). This is most likely the result of preferential volatilization of lithium during plasma spraying. However, the impurities and nonstoichiometry of the lithium ferrite do not pose serious problems in using it as an anode for cathodic protection.

The dissolution rates of ceramic anodes prepared from crushed lithium ferrite bars and from spray-dried powders were determined by calculating the overall anode dissolution rate from the proportions of lithium and iron in the test solutions. It was found that plasma-sprayed, spray-dried powders gave lower dissolution rates than the coatings prepared from crushed bars. (Only the results for the spray-dried powders are presented here.)

Chemical analysis (see Table 2) was used to calculate mole fractions of lithium and iron in the anodes. Table 3 shows the dissolution rates of the ceramic anodes at various current loadings of 3.5 percent sodium chloride in distilled water and in tap water. Figure 2 illustrates the effect of varying current density on the anode dissolution rates, based on both lithium and iron dissolution rates for tap water and saltwater. Error bars are shown for points where more than one test was performed; these error bars reflect the accuracy of all the data presented. It appears that in short-term tests, the anode dissolution rates are maximum at currents of 20 mA (≈ 20 A/m^2) in both saltwater and tap water. The anode dissolution rates approach unacceptable limits at 20 mA in saltwater (111 g per ampere year), and the dissolution rate of the ceramic anodes decreases with increasing current density at total currents above 20 mA. This is very desirable, because it suggests that

Table 2

Chemical Analysis of Plasma-Sprayed Lithium Ferrite
on Titanium Substrate*

	Weight Percent	Mole Percent
Iron	69.22	40.49
Manganese	3.05	1.81
Lithium**	1.46	6.87
Zinc	1.43	0.72
Magnesium	0.008	1.08×10^{-2}
Calcium	0.27	0.22
Nickel	0.082	0.46
Oxygen (estimated)	24.375	49.77
Others (Approximate):		
Titanium	0.02	1.36×10^{-2}
Aluminum	0.01	1.21×10^{-2}
Boron	0.01	3.02×10^{-2}
Silicon	0.0005	5.82×10^{-3}
Chromium	0.003	1.88×10^{-3}
Copper	0.003	1.54×10^{-3}
Molybdenum	0.003	1.02×10^{-3}
Cobalt	0.001	5.54×10^{-4}
Silver	0.0005	1.51×10^{-4}
Sod. - Questionable, If Present	0.05	≈ 0

*Analysis performed by x-ray fluorescence on a fused sample of plasma-sprayed
 lithium ferrite.
**Lithium analysis performed by atomic absorption.

the ceramic anodes may be well-suited to operation at very high current
densities, thus reducing size requirements.

The dissolution rate of lithium ferrite, found by weight loss meas-
urements, was only 1.7 g per ampere year at a current loading of 2000
A/m^2 during the 2-month, long-term test. This also shows that the dis-
solution rate of the oxide decreases with increasing current density.

Laboratory Tests on Sintered Anodes

Table 4 shows the results of the anodic polarization tests. Of the
solid ceramics investigated, manganese zinc ferrite and manganese fer-
rite appear to have the highest potential as feasible anode materials.
Sintered anodes made of barium and strontium titanate and titanium diox-
ide failed the anodic polarization tests because of the formation of an
insulating oxide during polarization. After anodically polarizing the

Table 3

Dissolution Rates of Lithium Ferrite Plasma-Sprayed Ceramic Anodes in Tap Water and 3.5 Percent Sodium Chloride in Distilled Water

No. of Tests	Solution	Current Density (A/m^2)	Total Current (mA)	Fe/Li Ratio in Solution	Anode Dissolution Rate Based on Fe (g/A-yr)	Anode Dissolution Rate Based on Li (g/A-yr)	Average Dissolution Rate Based on Average Li and Fe Dissolution Rates (g/A-yr)
1	Tap Water	15	15	14.7	3.9	11.6	7.8
3	Tap Water	20	20	43.7±4.6	17.6±5.0	20.8±4.0	19.2±9.0
1	Tap Water	44	44	47.1	8.01	8.1	8.05
3	Tap Water	100	100	84.3±8.4	6.9±1.2	16.7±4.81	11.8±6.0
1	Salt Water	15	15	41	13.26	15.1	14.2
3	Salt Water	20	20	9.7±0.3	110.7±22.5	545.7±126.8	328±150
3	Salt Water	100	100	151.2±24.7	21.72±0.9	29.3±7.7	22.5±8.6
1	Salt Water	200	200	20.0	0.46	1.09	1.55

Table 4

Results of Preliminary Pass-Fail Anodic Polarization Tests

Material	Performance	Typical Method of Failure (if applicable)
$BaTi_3$	Fail	Sample Spalled
$CaTiO_3$	Fail	Sample Stopped Passing Current
$SrTiO_3$	Fail	Sample Stopped Passing Current
$MnFe_2O_4$	Pass	N/A
$Mn_{.8}Zn_{.2}Fe_2O_4$	Pass	N/A
TiO_2	Fail	Sample Spalled
$Mn_{.5}Zn_{.5}Fe_2O_4$	Pass	N/A

samples for 100 to 200 hours in 3.5 percent sodium chloride in distilled water, dissolution rates for the materials which passed the initial test were calculated from the concentration of individual composition elements found in solution. It was found that the manganese zinc ferrite and manganese ferrite had similar dissolution rates, ranging from .384 to .331 g per ampere year (Table 5).

It is apparent that the hard sintered manganese and manganese zinc ferrites tested performed significantly better under comparable situations than the lithium ferrite. Further investigation will help determine if this is due to material or processing properties.

214
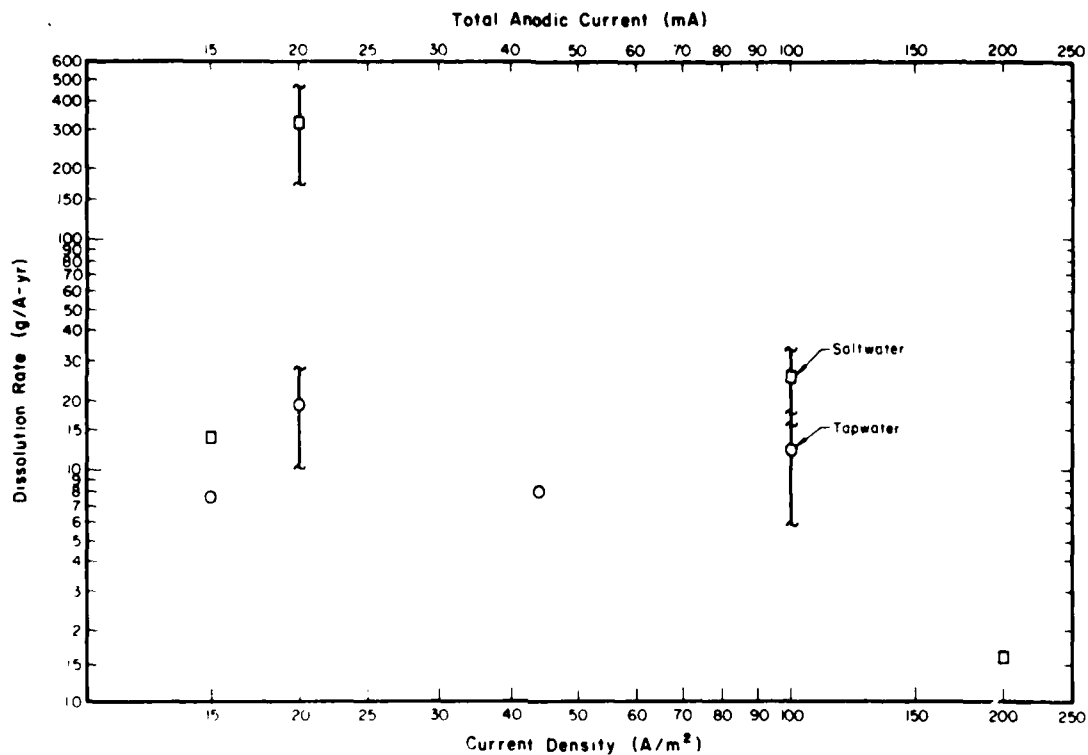
Figure 2. The effect of varying current density on the dissolution rate of lithium ferrite coated titanium anodes in tap water and in 3.5 percent sodium chloride in distilled water.

Field Tests

The average current flowing through the test ceramic anodes in the water storage tanks and in the fresh water canal gates was found to be 20 mA at 5 V applied potential. The resistivity of the water was between 3000 and 3500 ohm-cm.

The average current passing through the ceramic anode installed near an underground coated pipe in soil of 5000 ohm-cm resistivity was 20 mA at 5 V. Higher currents can be obtained by increasing the rectifier voltage or by using media of lower resistivity. This shows that one ceramic anode can protect 9 m^2 of coated steel (assumes 10 percent base area).

215

Table 5

Dissolution Rates of Sintered Ferrites in 3.5 Percent Sodium Chloride in Distilled Water

Number of Tests	Material	Current Density (A/m^2)	Anode Dissolution Rate (g/A-yr) Based on:			Average Anode Dissolution Rate (g/A-yr)
			Fe	Mn	Zn	
1	$MnFe_2O_4$	250	*1.85	*4.33	–	*3.09
1	$MnFe_2O_4$	1000	.736	.032	–	.384
1	$Mn_{.8}Zn_{.2}Fe_2O_4$	1000	.462	.203	*5.26	.332
1	$Mn_{.8}Zn_{.2}Fe_2O_4$	100	.529	.134	*45.3	.331
1	$Mn_{.5}Zn_{.5}Fe_2O_4$	100	*4.5	*27.5	*231	*87.6

*Based on worst-case calculations.

Advantages of Ceramic Anodes

On the basis of the information provided by these tests, ceramic anodes appear to have the following advantages over the presently used anodes:

a. The substrate metals used for ceramic anodes, such as titanium or niobium, can be easily fabricated in any shape or form. The substrates also offer a good strength-to-weight ratio.

b. The consumption rate of ceramics is about one one-hundredth that of graphite or HSCBCI. Therefore, smaller anodes can be used which are less vulnerable to damage. Preliminary evidence suggests that the dissolution rate decreases with increasing current density for total anode currents of 20 to 100 mA. This suggests that small anodes with high current loadings may perform best.

c. Small anodes can be manufactured in a factory, so less fabrication will be required in the field. This increases the reliability of the cathodic protection system.

d. The smaller size of the anodes makes their replacement much easier.

e. Lightweight metal-ceramic anodes can be installed using plastic supports. This reduces installation problems which arise from the difficulty of providing electrical isolation of metallic supports used for the heavier button-type anodes which can weigh more than 60 lb.

f. Ceramic materials are naturally resistant to abrasion; the inherent problem of ceramics brittleness can be overcome with appropriate design modifications.

Conclusions and Recommendations

This research has produced the following conclusions:

a. Anodes made with lithium ferrite coatings on a pure titanium or niobium substrate exhibit good electrical conductivities and long-range stability and therefore have proved to have excellent applicability to cathodic protection systems.

b. The dissolution rate of the ferrite field anodes during short-term tests at a current loading of 20 A/m^2 is about 10 g per ampere year in fresh water (one one-hundredth of the currently used HSCBCI and graphite anodes).

c. The dissolution rate of plasma-sprayed lithium ferrite in 3.5 percent sodium chloride in distilled water was 1.7 g per ampere year at 2000 A/m^2 over a 2-month period.

d. Lithium ferrite anodes exhibit decreasing dissolution rates with increasing current densities when the current densities are above 20 A/m^2.

e. A unique button-shaped anode was developed which can be easily installed on underground pipes and submersed structures. Structures in water, such as waterway lock gates and elevated water storage tanks, can be protected without dewatering the structure. These anodes are tough and are small enough to be recessed in locations where damage from debris and ice is less likely.

f. The preliminary investigation of sintered manganese and manganese zinc ferrites suggests that these materials exhibit a dissolution rate of less than 1 g per ampere year.

The following recommendations are made concerning the further development of these anodes:

a. Investigate and develop improved materials and designs for ceramic anodes for use in specialized environments encountered in hydraulic structures.

b. Continue field and laboratory testing of semiconducting ceramics to provide a full understanding of the short- and long-range behavior of ceramic-coated anodes.

c. Investigate the mechanisms of dissolution and the effect of structure and composition on the electrochemical stability of the ceramic semiconductors now being used as anodes in impressed current cathodic protection systems. Use the results of the mechanistic studies to develop new ceramics for anodes that offer improved properties. The development of materials that exhibit substantially improved dissolution characteristics hinges on an understanding of the factors that affect dissolution. Unlike the extensive base of knowledge of the electrochemistry of metals, the electrochemistry of ceramic semiconductor anodes for cathodic protection is not well understood.

Acknowledgements

References

(1) S. Wakabayashi and T. Aoki, "Characteristics of Ferrite Electrodes," Journal De Physique, Vol. 4 (1977), p. C1.

(2) R. Itai and H. Kanai, U.S. Patent No. 3,850,701, "Anode Coated with Magnetite and the Manufacture Thereof" (November 1974).

(3) T. Fujii, T. Kodama, H. Baba, and S. Kitahara, "Anodic Behavior of Ferrite Coated Titanium Electrodes," Boskoko Gijutsul (Corrosion Engineering) Vol. 29, No. 4 (1980), pp. 180-184.

THE PARALYTIC EFFECT OF BOTULINAL NEUROTOXINS
AND POTENTIAL THERAPEUTIC MODALITIES

LAWRENCE C. SELLIN, Ph. D.
U.S. ARMY MEDICAL RESEARCH INSTITUTE OF INFECTIOUS DISEASES
FORT DETRICK, FREDERICK, MD 21701

Botulism results from the action of the most potent of the known
biological or chemical toxins. Botulinum toxin (BoTx), a protein neuro-
toxin (m.w. about 150,000), has a human lethal dose estimated to be as
low as 1 ng/kg (1). The primary site of action of BoTx is the cholin-
ergic nerve terminal where it inhibits the release of the neurotrans-
mitter, acetylcholine. Its action is most prominent in cranial nerves
and at the neuromuscular junction. The cause of death is usually
respiratory paralysis, due to blockade of transmitter release from the
phrenic nerve to the diaphragm muscle (see reviews 7,8).

Botulinal neurotoxins are found in seven immunologically distinct
types, A-G. Four types have been identified in cases of human botulism
and they occur at a frequency of A>B>E>F (see review, 8). Although our
understanding of botulism and the action of BoTx has increased, two areas
have not been adequately addressed. These are: studies comparing the
specific effects of the different types and new treatment modalities for
botulism.

A study comparing the clinical features of types A and B botulism
suggested that type A was more severe than type B (3). Patients with
type A botulism saw physicians earlier, were more likely to require
respiratory support, and were hospitalized longer than those with type B
disease. Beyond these observations there is little or no information
concerning the differences in the action of the various BoTx types.

The traditional treatment for botulism has been administration of
antitoxin and respiratory support. However, there are two problems with
antitoxin therapy. First, the only licensed antisera are derived from
horse serum which produce side reactions in about 20% of recipients and
anaphylaxis in 3% (4). Second, although antitoxin is a useful
therapeutic agent, it is only effective on circulating toxin (14).

219

Antitoxin does not reduce the paralysis once the toxin exerts its effect. Therefore, agents which increase transmitter release are likely candidates for relieving the paralysis caused by BoTx. Aminopyridines, particularly 3,4-diaminopyridine (3,4-DAP) are known to stimulate neural transmitter release in normal and toxin-treated preparations. (5).

The experiments described in this paper were designed to compare the actions of the different types of BoTx and evaluate the efficacy of 3,4-DAP as an antagonist to its paralytic action.

METHODS

Experiments were performed <u>in situ</u> and <u>in vitro</u> on the extensor digitorum longus (edl) muscle of male Wistar rats (100 to 200 g). A single 0.25 ml bolus of BoTx in gel phosphate buffer (13) was injected subcutaneously into the anterolateral region of the right hind leg, superficial to the distal part of the tibialis anterior muscle. At 1, 3, 7 or 10 days after injection, the edl nerve-muscle preparation was examined for alterations in muscle mechanical properties (<u>in situ</u>) or electrophysiological (<u>in vitro</u>) properties. BoTx toxicity was determined by mouse bioassay (6,13) and expressed as number of mouse i.p. LD_{50} (the amount of protein toxin capable of producing 50% lethality). In some experiments BoTx was treated with 0.2 M ethylacetimidate in 28 mM Na_2HPO_4 buffer solution for 30-40 minutes to "amidinate" lysine residues (15).

Single Twitch Tension.

Rats were anesthetized with α-chloralose (75 mg/kg, i.p.) (CalBiochem, La Jolla, CA) and methoxyflurane (Penthrane, Abbott Laboratories, North Chicago, IL) in preparation for measurement of single twitch tension (muscle contraction). The tendons of insertion for the edl were isolated together with 1 cm of the deep peroneal nerve. The sciatic nerve was cut to prevent retrograde stimulation. The surrounding muscles and vasculature were not damaged although the tendon of insertion for the tibialis anterior muscle and the transverse ligament were cut. At this time, methoxyflurane anesthesia was discontinued, but the rats were still fully anesthetized by the effects of the α-chloralose. The tendons of insertion of the edl were connected, via a gold chain, to a force transducer (FT 0.03) and tension was recorded on a Model 5B polygraph (Grass Instruments, Quincy, MA).The deep peroneal nerve was stimulated with short duration (0.2 to 0.5 msec) and maximal amplitude (4 to 7 V) square pulses using a Grass S88 stimulator. Resting length was adjusted to achieve maximal single twitch tension. Exposed areas were kept moist with sterile saline irrigation solution and temperature was maintained using a heating pad and heat lamp.

Electrophysiology.

The edl nerve-muscle preparations were excised during methoxyflurane anesthesia, under a continuous flow of oxygenated (95% O_2, 5% CO_2) Kreb-Ringer solution (composition in mM: NaCl, 132; KCl, 5; $MgCl_2$, 2; $CaCl_2$, 2; Na_2HPO_4, 1; $NaHCO_3$, 18; dextrose, 11; pH, 7.2 to 7.3). The muscles, pinned through their tendons and stretched to approximately their resting lengths, were placed in a 12-ml chamber (30 ± 1°C) and suffused at a rate of 3 to 4 ml/min. After a 20 to 30 min equilibration period, transmitter release was examined by recording end-plate potentials (e.p.p.s). Glass microelectrodes, filled with 3 M KCl, having tip resistance of 5-10 x 10^6 Ω were used. E.p.p.s were recorded following nerve stimulation (10 to 50 μsec, 2 to 4 V) via a platinum wire-glass capillary suction electrode. (For details see: 9,10,12).

RESULTS

The fundamental differences in the neuromuscular effects of type A and types B, E and F were in terms of potency, duration of action and response to 3,4-DAP. In general, type A was more potent, paralyzed for a longer period of time and type A paralysis was more sensitive to the paralysis-reversing (stimulatory) effect of 3,4-DAP than types B, E or F.

Subcutaneous injection of BoTx types A, B, E or F in the lower hindlimb of the rat produced unilateral paralysis within six to twelve hours. In all cases, larger doses of types B, E and F (>200 LD_{50}) were required to produce a paralysis comparible to that produced by relatively low doses of type A (5 LD_{50}). Despite the requirement for larger doses, rats treated with types B, E or F recovered at a much faster rate (5 days) than type A-treated animals (>14 days). Rapid recovery was also observed in animals treated with very high (barely sublethal) doses (≈5000 LD_{50}) of types B,E or F. These observations are illustrated in a comparison of muscles treated with types A and F BoTx (Figure 1).

Precise measurements of transmitter release can be obtained in vitro by intracellular recording with glass microelectrodes. In BoTx-paralyzed muscles extremely small amounts of transmitter (acetylcholine) are released after nerve stimulation. The postsynaptic electrical responses (end-plate potentials) produced by the released transmitter can be recorded with the microelectrode inserted in the muscle fiber at the neuromuscular junction.

Addition of 3,4-DAP increased the amplitude and frequency of end-plate potentials (e.p.p.s) in muscles treated with types A, B, E or F BoTx. However, it did so with different efficacies. That is, 3,4-DAP was somewhat less effective in reversing the paralysis produced by types B, E and F than for muscles treated with type A BoTx (Figure 2). Addition of 100 µM 3,4-DAP caused all type A-treated muscles to twitch after nerve stimulation, regardless of amount of toxin administered. A similar response was not seen in all muscles treated with types B, E or F BoTx.

SINGLE TWITCH TENSION

DAYS AFTER BoTx (S.C.) INJECTION
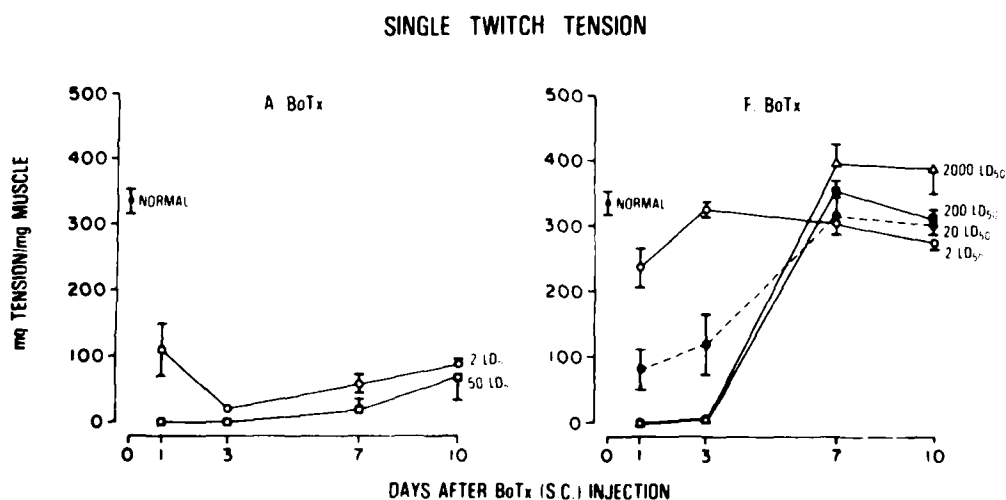
Figure 1. Nerve stimulus evoked single twitch tension (muscle contraction) of the edl muscle recorded in situ from anesthetized rats and plotted against days after BoTx injection. Each point, normalized for muscle weight, is the mean ± SEM for at least 10 muscles (11).

222
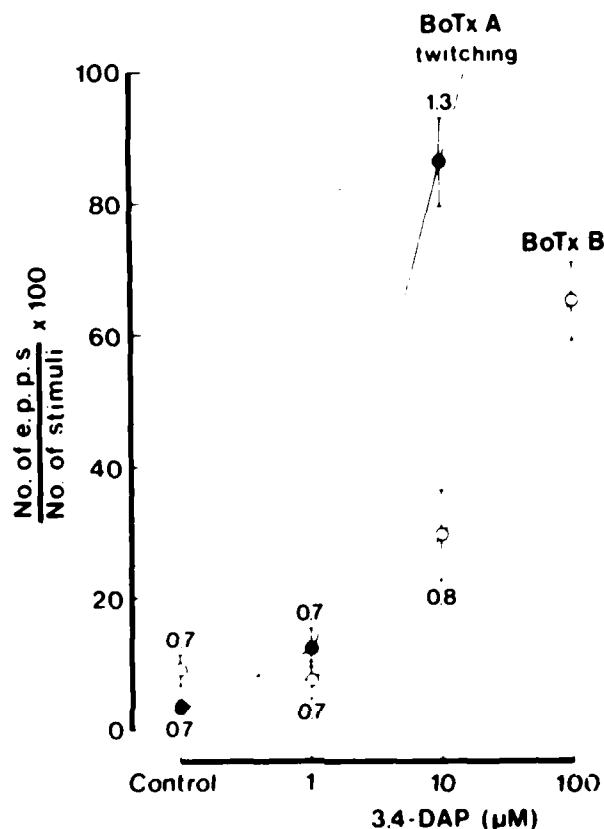
Figure 2. Effect of various concentrations of 3,4-DAP on
            neuromuscular block produced by BoTx A and B 3 days after
            poisoning. The graph shows the number of nerve stimuli
            causing transmitter release (e.p.p.s) as a percentage of
            the total number of stimuli at 1 Hz. Each point
            represents mean ± SE from at least 8 fibres in 2
            muscles. The numbers above or below the points are the
            mean amplitude in mV of recorded e.p.p.s. Note that with
            100 μM 3,4-DAP, BoTx type A-poisoned muscles twitched
            (12).

An explanation for the reduced effect of 3,4-DAP in muscles treated
with BoTx types B, E or F was the presence of asynchronous transmitter
release as shown in Figure 3. No such effect was observed in type A-
treated muscles.

BoTx A   BoTx B

Control

+10 µM 3,4-DAP

+100 µM 3,4-DAP

1mV

4 ms
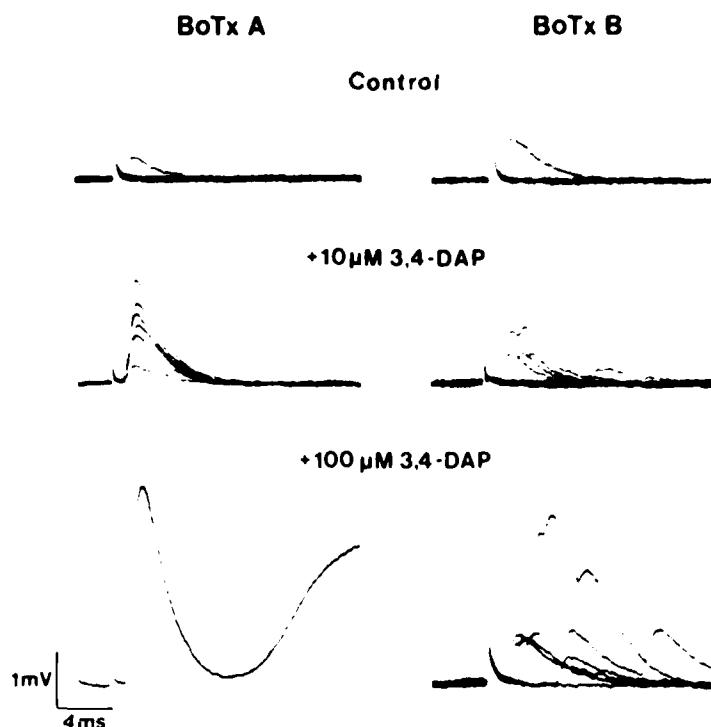
Figure 3.   Examples of e.p.p.s (superimposed tracings) in type A and
B poisoned muscles 3 days after injection and the effects
of 10 and 100 µM 3,4-DAP. Stimulation frequency was 1
Hz. Note that in type A poisoned muscles. 3,4-DAP
enhances synchronous transmitter release and at 100 µM
causes the muscle to twitch (lowest record). In contrast
3,4-DAP causes asynchronous transmitter release and no
twitches in a type B poisoned muscle (12).

Potentiation of the effect produced by 3,4-DAP in BoTx-treated
muscles was observed with combined administration of 3,4-DAP and the
acetylcholinesterase inhibitor neostigmine (or pyridostigmine) or 3,4-DAP
and theophylline.

In an effort to analyze the basis for the observed functional
difference between the BoTx types, chemical modification by "amidination"
was attempted (2,15). Treatment of type F toxin with ethyl acetimidate
increased its toxicity by 1.5-8X after i.p. injection into mice. This

procedure also increased the duration of paralysis of the edl after s.c. injection in rats (Figure 4). Presumably, the reaction of imidoesters with protein amino groups (most likely lysine residues) forms amidines which are stronger bases than their parent amines. This factor may play an important role in determining the relative potencies of the toxins. However, preliminary experiments indicate that treatment with ethylacetimidate does not alter the toxin's antitoxin binding site.

SINGLE TWITCH TENSION
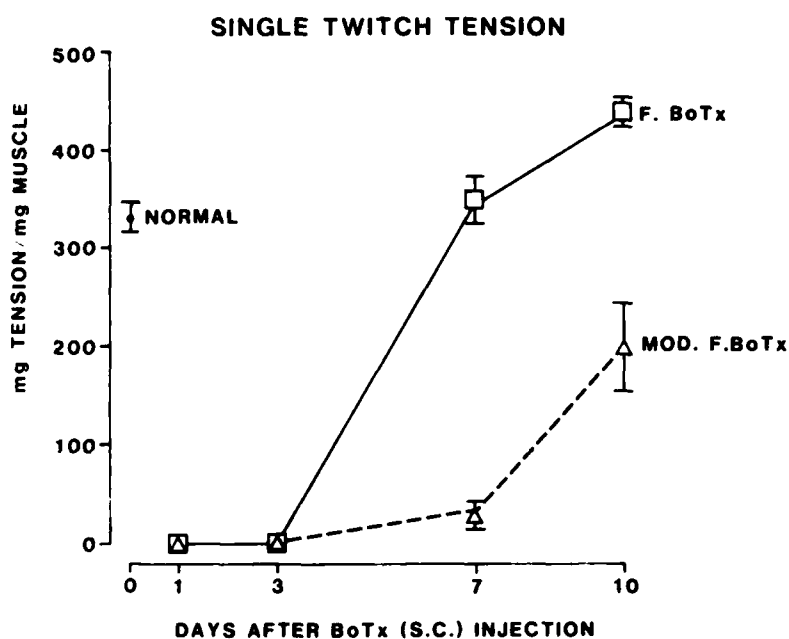
Figure 4. Nerve stimulus-evoked single twitch tension (muscle contraction) of the edl muscle recorded in situ from anesthetized rat and plotted against days after BoTx injection. One group was given native F BoTx (1000 LD_{50}) while the other was given the same dose of F BoTx but modified with ethylacetimidate prior to injection. Each point, normalized for muscle weight, is the ± SEM for at least 4 muscles (11).

225

DISCUSSION

The results demonstrate that type A BoTx is more potent, paralyzes for a longer period of time and type A paralysis is more sensitive to the paralysis-reversing action of 3,4-DAP than types B, E or F. The observed differences in the effects of the various types of BoTx supports, in part, the differences seen in clinical cases of types A and B botulism (3). The specific structural differences between the protein neurotoxins which determine their pharmacological activities have not been identified. However, the present data with ethylacetimidate-modified toxin implicate lysine residues as important components affecting toxicity. It is interesting to note that this is the first demonstration of a chemical modification which increases the activity of one type of BoTx. However, this same modification does not appear to alter the toxin's antitoxin binding site. Increases in ribonuclease A activity by treatment with a bifunctional imidoester has also been demonstrated (2). In contrast, the amidination reaction did not alter antibody binding capacity of bovine serum albumin or rabbit antibody to benzene-arsonic acid (15).

All BoTx types reduced nerve-evoked transmitter release as measured with microelectrodes in vitro. However, 3,4-DAP was more effective in restoring neuromuscular transmission in type A-treated muscles than in muscles receiving types B, E or F. One explanation may be the asynchronous release observed in muscles treated with B,E or F BoTx. Although 3,4-DAP may not prove to be an ideal antidote for all cases of botulism, it may be useful in delaying the onset of respiratory paralysis, "weaning" patients off respirators or providing the basis for the development of additional compounds or combinations of compounds. Furthermore, combinations of compounds such as 3,4-DAP, theophylline and/or neostigmine-pyridostigmine may provide a "generic" therapy for a variety of neuromuscular toxins.

ACKNOWLEDGEMENTS

REFERENCES

1. Gill, D. M. 1982. Bacterial toxins: a table of lethal amounts. Microbiol. Rev. 46:86-94.

2. Hartman, F. C. and Wold, F. 1967. Cross-linking of bovine pancreatic ribonuclease A with dimethyl adipimidate. Biochemistry 6:2439-2448.

3. Hughes, J. M. et al. 1981. Clinical features of types A and B foodborne botulism. Ann. Intern. Med. 95:442-445.

4. Merson, M. H. et al. 1974. Current trends in botulism in the United States. JAMA 229:1305-1308.

5. Molgo, J. Lundh, H. and Thesleff, S. 1980. Potency of 3,4-diaminopyridine and 4-aminopyridine on mammalian neuromuscular transmission and effect of pH changes. Eur. J. Pharmacol. 61:25-34.

6. Reed. L. J. and Muench, H. 1983. A simple method of estimating fifty percent end-points. Am. J. Hyg. 27:493-497.

7. Sellin, L. C. 1981. The action of botulinum toxin at the neuromuscular junction. Medical Biology 59:11-20.

8. Sellin, L. C. 1984. Botulism - an update. Military Medicine 149:12-16.

9. Sellin, L. C. and Thesleff, S. 1981. Pre- and postsynaptic actions of botulinum toxin at the rat neuromuscular junction. J. Physiol. 317:487-495.

10. Sellin, L. C., Kauffman, J. A. and Das Gupta, B. R. 1983. Comparison of the effects of botulinum neurotoxin types A and E at the rat neuromuscular junction. Medical Biology 61:120-125.

11. Sellin, L. C., Kauffman, J. A., Way, Jr., J.F. and Siegel, L. S. 1983. Comparison of the action of types A and F botulinum toxin at the rat neuromuscular junction. Neuroscience Abstracts 9:1103.

12. Sellin, L. C., Thesleff, S. and Das Gupta, B. R. 1983. Different effects of types A and B botulinum toxin on transmitter release at the rat neuromuscular junction. Acta Physiol. Scand. 119:127-133.

SELLIN

13. Siegel, L. S. and Metzger, J. F. 1979. Toxin production by *Clostridium botulinum* type A under various fermentation conditions. Appl. Environ. Microbiol. 38:606-611.

14. Simpson, L. L. 1980. Kinetic studies of the interaction between botulinum toxin type A and cholinergic neuromuscular junction. J. Pharmacol. Exp. Therapeut. 212:16-21.

15. Wofsy, L. and Singer, S. J. 1963. Effects of the amidination reaction on antibody activity and on the physical properties of some proteins. Biochemistry 2:104-116.

DEVELOPMENT OF POROPLASTIC® AS A WOUND DRESSING MATERIAL
FORMULATED TO SLOWLY RELEASE GENTAMICIN-SULFATE (U)

JEAN A. SETTERSTROM, Ph.D.*
JACK W. VINCENT, COL, DC
LARRY D. NICHOLS, Ph.D.†
U.S. ARMY INSTITUTE OF DENTAL RESEARCH
WASHINGTON, DC   20307
AND
MOLECULON RESEARCH CORPORATION†
CAMBRIDGE, MASSACHUSETTS   02142

INTRODUCTION

Avulsions, superficial and full-thickness burns, abrasions, and lacerations are commonly incurred in combat. Current wound care generally includes application of non-occlusive gauze in combination with applied pressure for hemorrhage control. Such dressings may provide a partial physical barrier from outside contamination and may retard bleeding, however, infection is not reduced and wound healing is not enhanced. Surface desiccation and adherence of the dressing to the wound are problems that retard the healing process. If the wound is kept moist, allowing no eschar formation, and is adequately oxygenated, epithelization will occur at a maximum rate. (1)

An ideal wound dressing should be impermeable to bacteria; control water vapor transmission to keep wounds moist, but not wet; allow oxygen exchange; conform, but not adhere to the wound; prevent infection; and control capillary and small vessel bleeding. The dressing should be light-weight, with stability for long periods under adverse storage conditions. It should be applicable to all partial and full-thickness skin loss injuries via a simple, single-step procedure. Traumatic wounds suffered in the field are generally heavily contaminated with foreign material, therefore, over a period of time the wound dressing should gradually deliver broad spectrum antimicrobial agents that are nontoxic to the injured tissue. Once applied, the bandage should remain tightly adherent to the intact skin and preferably offer a watertight seal. It should remain supple, resilient, and allow normal wound contraction.

In the present study, we have capitalized on the properties of a

229

unique, ultramicroporous, homogenous membrane (2). POROPLASTIC® is a hydrogel in which the liquid content can be varied from 70-98% water by weight. It consists of two interpenetrating and largely independent phases. One component is a strong, thermally stable cellulose triacetate resin and the other component can be almost any liquid or solution. This characteristic allows POROPLASTIC® to be impregnated with alcohols, esters, liquid monomers, hydrocarbons, salt or dye solutions or a variety of medicaments. POROPLASTIC® is strong, flexible, nonionic, and isotropic with excellent pH stability.

The pore size, flux and water content of POROPLASTIC® can be engineered. By changing the thickness or the water content, hydraulic permeability can be controlled. The pore diameters of POROPLASTIC® range from 0.0014μ for films containing 70% water by weight (MA-70) to > 0.02μ for films containing 97.5 % (MA-97.5). The less water in the dressing (MA-70), the stronger the film (tensile strength, 1300 psi). The small pore size assures bacterial impermeability, but allows water vapor transmission. Physical characteristics of POROPLASTIC® make it extremely useful as a vehicle for the controlled sustained release of medicaments. In this study we have evaluated POROPLASTIC® impregnated with gentamicin sulfate for effectiveness in eliminating Staphylococcus aureus from both acute (contaminated) and established wound infections. Work is proceeding to evaluate the effectiveness of POROPLASTIC® in the enhancement of wound healing and to evaluate dressings that release dual antibiotics to assure a broader spectrum of antimicrobial activity.

MATERIALS AND METHODS

POROPLASTIC® Membranes

POROPLASTIC® wound dressings are made from an acetic acid-cellulose triacetate dope that is blade-casted on an inert surface and cured in water (2). The water cure removes all traces of acetic acid, yielding, after separation from the casting surface, a durable, transparent, water-impregnated cellulose triacetate membrane. Impregnation of POROPLASTIC® with gentamicin sulfate is accomplished by simple immersion into the desired concentration of the drug. Following absorption of drug, the membranes are immersed in anhydrous isopropyl alcohol to replace the aqueous component and cause deposition of gentamicin sulfate within the membrane. In a final step, the alcohol is exchanged for an inert, hydrophobic liquid.

POROPLASTIC® membranes produced for evaluation in this study ranged in thickness from 13-23 mil. They were impregnated with amounts of gentamicin sulfate ranging from 4.3-8.4 wt%. Parallel production of genta-

micin-free POROPLASTIC® provided wound dressings for untreated control animals. All POROPLASTIC® wound dressings were hermetically sealed and sterilized by exposure to 2.5 Mrad of gamma radiation (Neutron Products, Inc., Dickerson, Maryland) before in vivo evaluation.

In Vitro Evaluation

Gentamicin-POROPLASTIC® wound dressings were evaluated in vitro to determine the total quantity of gentamicin sulfate impregnated within the dressing and establish the rate of drug release into a receiving fluid at 37°C. Small weighed samples of gentamicin-POROPLASTIC® were put in tubes containing known excess quantities of alcohol to allow complete displacement of isopropyl myristate from the dressing. Following the addition of known quantities of distilled water to each tube to solubilize the readily diffusible gentamicin sulfate, the aqueous solutions were assayed for drug by the method of Shriner et al. (3).

The rate of gentamicin sulfate release from POROPLASTIC® was determined by exposing preweighed samples of POROPLASTIC® to known quantitites of 0.02M potassium phosphate buffered saline (0.1M, pH 7.4). The receiving fluid was assayed at timed intervals for gentamicin sulfate concentration. The per cent cumulative release (μg/cm^2) was plotted as a function of time (hours). In vitro release curves were determined for individual lots of dressing material evaluated in vivo.

In Vivo Evaluation

To evaluate both antibiotic-loaded and unloaded POROPLASTIC® in vivo, full thickness wounds approximately 676 mm^2 were excised from the preshaven interscapular area of anesthetized guinea pigs. Each wound was then inoculated with 2-4 x 10^9 Staphylococcus aureus (ATCC 12600), and covered with a POROPLASTIC® dressing (961 mm^2). Equal numbers of untreated and treated animals were included in each experiment. Dressings were held in place by adhering plastic squares of clear polyethylene (1600 mm^2, 5 mil) to the intact skin with cyanoacrylate adhesive, either Histoacryl®blue tissue adhesive, (B. Braun Melsungen AG, Federal Republic of Germany) or Wonder Bond® (Borden, Inc., Columbus Ohio). Both acute (contaminated) and reinfected established infections (wounds with >10^6 S. aureus/cm^2 at three days post initial contamination) were evaluated.

The viable count of S. aureus/cm^2 of wound surface was determined on the third day following dressing of the wound with POROPLASTIC®. Following anesthesia, a sterile plastic cylindrical chamber (2.54 cm^2) was held firmly to the wound surface. A 1 ml solution 0.1% Triton X-100 in 0.075M phosphate buffer, pH 7.9 was transferred by sterile pipette into the chamber and the area was scrubbed with moderate pressure for 20 seconds

using a sterile Teflon "policeman". The wash fluid was aspirated, re-
placed with 1 ml of fresh scrub solution and the scrub was repeated. To
prevent bacterial aggregation and thoroughly clean the area, two addi-
tional scrubs each with 1 ml of 0.05% Triton X-100 in 0.0375M phosphate
buffer were performed and all of the wash solution (4 ml) was pooled for
assay. The number of bacteria in the inoculum and wash solutions from
each wound were determined by a mechanical-bioassay method using a spiral
plater and laser bacterial colony counter (Spiral System Instruments,
Inc., Bethesda, Maryland). The number of S. aureus/cm^2 of wound surface
of both treated and untreated wounds was calculated.

Animals initially treated with unloaded POROPLASTIC® were subsequent-
ly used to evaluate the efficacy of wound dressings on established
wound infections. These animals were subdivided into two groups on day
three. Following the scrub procedure, both groups were reinfected with
2-4 x 10^9 S. aureus. Group one received unloaded POROPLASTIC®, while
group two received gentamicin POROPLASTIC®. The wound dressings were
covered with polyethylene films and secured with cyonoacrylate adhesive.

Serum was obtained from randomly selected animals. Sera and wash
solutions then were assayed for gentamicin sulfate concentration by
microbial bioassay (4). The microbial bioassay for gentamicin sulfate
was performed on neomycin assay agar (BBL Microbiology Systems,
Cockeysville, Maryland). Gentamicin standards for the bioassay were
prepared from gentamicin sulfate (Schering Corporation, Kenilworth, New
Jersey) by dissolving crystalline drug weighed on a Cahn Electrobalance,
(Model 29, Cerritos, California) into a known amount of sterile distilled
water. Serial dilutions were then made to obtain known standards. Un-
known specimens and standards were then placed on sensitivity discs and
positioned on duplicate plates containing agar seeded with Staphylococcus
epidermidis (ATCC 27626). Plates were incubated at 37°C for 18 hours.
Zones of inhibition were measured to the nearest 0.1 mm.

Cytotoxicity was evaluated by placing small pieces of sterile
POROPLASTIC® (antibiotic-loaded or unloaded) on Vero cells ATCC CCL 81
grown in Dulbecco's Modified Eagle Media (DMEM) supplemented with 10%
fetal calf serum, 100 mM L-glutamine, 1.0% Na-pyruvate, 1.0% nonessential
amino acids, 1.0% NCTC 109 (National Cancer Tissue Culture Media),
penicillin (100 units/ml), streptomycin (100 µg/ml), and Fungizone
(0.25 µg/ml). A suspension of 1.0 x 10^6 viable Vero cells/ml was dis-
pensed (400 µl per well) into 24-well Costar tissue culture plates.
Following aseptic placement of the POROPLASTIC® membrane into individual
wells, the plates were incubated at 37°C in 7% CO_2 at 100% humidity.
They were observed for effect on cell growth at one, three, and four
days.

FIGURE 1
In Vitro Release from Lot No. 250-086-C
7% Gentamicin

FIGURE 2
Gentamicin Release Rate vs. Loading
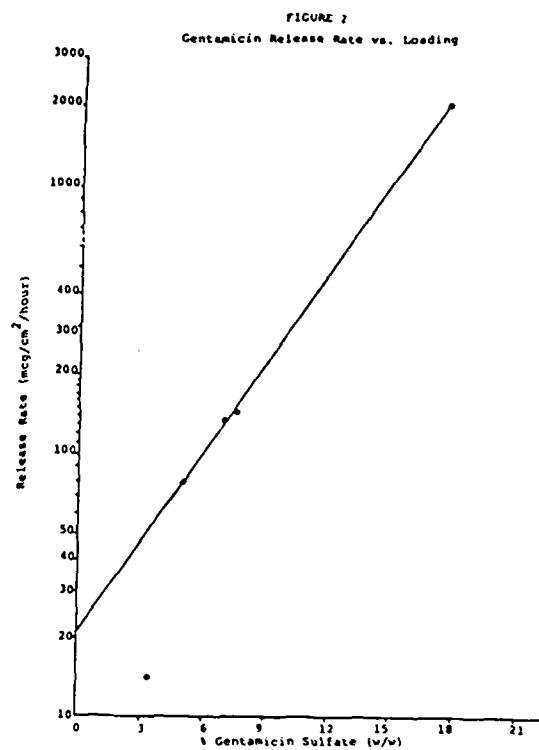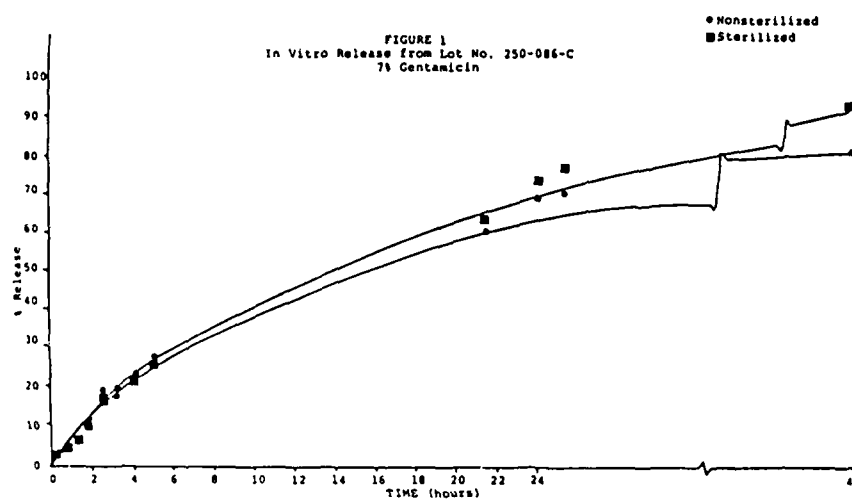
RESULTS

Typical in vitro release profiles of gentamicin sulfate from sterilized and unsterilized POROPLASTIC® are compared in Figure 1. No significant radiation effect was noted on the in vitro release rate profiles. The dependence of release rates on the POROPLASTIC® drug load is shown in Figure 2. These data are presented as the release rate (μg/cm^2/hour) of gentamicin sulfate as a function of time. As shown, the rate of release of gentamicin from POROPLASTIC® in vitro is highly dependent on loading and approximates zero order kinetics.

The number of bacteria present on acutely infected (contaminated) wounds three days post-treatment either unloaded POROPLASTIC® and with gentamicin-POROLASTIC® are shown i. Table 1. Even though animals in this experiment removed their dressings within 24 hours of application, 83.3% of the wounds covered with gentamicin-POROPLASTIC® (6.0-6.5 wt%) were sterile when assayed at three days post-wounding, and 22.2% were sterile

TABLE 1. Number of Staphylococcus aureus/cm^2 Remaining on Wound Surfaces at Three Days Following Inoculation With 4.5 X 10^9 S. aureus/Wound.

Animal	Unloaded POROPLASTIC®	Animal	Gentamicin POROPLASTIC®	Gentamicin Content (wt%)
1	6.29 X 10^6	21	89	5.8
2	3.55 X 10^7	22	0	6.1
3	7.00 X 10^7	23	0	5.7
4	3.55 X 10^7	24	0	5.5
5	9.60 X 10^6	25	0	6.5
6	1.22 X 10^8	26	0	6.4
7	5.02 X 10^6	27	0	6.0
8	2.15 X 10^8	28	1.14 X 10^6	5.7
9	ND	29	1.41 X 10^4	5.7
10	3.93 X 10^7	30	15	4.7
11	5.98 X 10^7	31	4.25 X 10^5	5.1
12	5.98 X 10^7	32	4.25 X 10^4	5.7
13	7.36 X 10^7	33	8.75 X 10^5	5.1
14	1.44 X 10^8	34	12	4.9
15	1.55 X 10^7	35	3.08 X 10^4	4.0
16	9.9 X 10^7	36	4.44 X 10^4	4.5
17	6.02 X 10^7	37	94	4.3
18	2.00 X 10^7	38	23	6.3
19	4.80 X 10^7	39	0	6.2
20	6.02 X 10^7	40	3	5.9

ND: Not Done (Anesthesia Death)

when covered with dressings containing 5.0-5.9 wt%. However, all five wounds covered with dressings containing 4.0-4.9 wt% gentamicin remained contaminated although with lower quantities of $\underline{S}$. $\underline{aureus}$ than untreated controls.

To evaluate wound dressings on established infections, animals 1-20 (Table 1) were divided into two groups. Following the three-day scrub assay to determine the bacterial count (>10^6 $\underline{S}$. $\underline{aureus}$/cm^2), the wounds were reinoculated with 2.3×10^9 $\underline{S}$. $\underline{aureus}$. Bacterial counts obtained four days after the wounds were covered with unloaded POROPLASTIC® (even-numbered) or gentamicin-POROPLASTIC® dressings (6.0-8.4 wt%, odd-numbered) are shown in Table 2.

TABLE 2. Number of $\underline{Staphylococcus\ aureus}$/cm^2 Remaining on Wound Surface* At Four Days Following Treatment of Established Wound Infections.

Animal	Unloaded POROPLASTIC®	Animal	Gentamicin POROPLASTIC®	Gentamicin Content (wt%)
2	1.73×10^7	1	0	8.0
4	1.9×10^6	3	17	8.4
6	3.8×10^7	5	6	6.5
8	4.8×10^7	7	6.0×10^3	7.5
10	2.4×10^7	9	ND	6.1
12	2.1×10^7	11	0	6.2
14	1.9×10^7	13	ND	6.0
16	Contaminant	15	25	6.2
18	1.73×10^7	17	Contaminant	6.4
20	6.2×10^7	19	0	6.9

*4.5×10^9 $\underline{S}$. $\underline{aureus}$ inoculated on Day 1; 2.3×10^9 $\underline{S}$. $\underline{aureus}$ inoculated on Day 3.

ND: Not Done (Anesthesia Death)

Results indicated that a wound dressing containing 6.2 wt% (Tables 1, 2) gentamicin held securely in place should assure wound sterility. Such dressings were formulated and evaluated in a subsequent experiment using a stronger adhesive (Wonder Bond®) to assure wound coverage for the entire three days.

As expected, wounds contaminated with 2.3×10^9 S. aureus and treated by covering with gentamicin-POROPLASTIC® (6.2 wt% MA-92, 20±1 mil) were free of S. aureus at three days (Table 3). Wounds contaminated and covered with unloaded-POROPLASTIC® or with polyethylene alone all harbored $>10^6$ S. aureus/cm² (Table 3). Animals (61-80) with established three-day infections ($>10^6$ S. aureus cm²) were then divided into two groups for subsequent treatment. All animals (61-70) treated with gentamicin-POROPLASTIC® were completely free of S. aureus three days later. All animals (71-80) treated with unloaded POROPLASTIC® harbored $>10^6$ S. aureus/cm² (Table 4). Gentamicin sulfate concentrations detected in scrub solutions from the wounds of these animals are shown in Table 5.

Table 3. Number of Staphylococus aureus/cm² Remaining on Wound Surfaces Three Days Following Inoculation With 2.3×10^9 S. aureus/Wound.

Animal	Unloaded POROPLASTIC® (22 ± 2) mil)	Animal	Gentamicin POROPLASTIC®* (20 ± 1 mil)
61	8.9×10^6	41	0
62	2.7×10^8	42	0
63	7.9×10^6	43	0
64	3.2×10^8	44	0
65	4.8×10^8	45	ND
66	7.2×10^6	46	0
67	2.7×10^7	47	0
68	1.1×10^8	48	0
69	2.3×10^7	49	0
70	1.3×10^7	50	0
71	8.6×10^6	51	0
72	2.4×10^8	52	0
73	3.4×10^7	53	0
74	1.2×10^7	54	0
75	8.6×10^6	55	0
76	3.2×10^8	56	0
77	4.7×10^8	57	0
78	9.8×10^7	58	0
79	1.4×10^8	59	0
80	3.2×10^6	60	0

*Gentamicin-POROPLASTIC® (6.2 wt%)

ND: Not Done (Anesthesia Death)

Table 4. Number of <u>Staphylococcus aureus</u>/cm^2 Remaining on Wound Surfaces* Three Days Following Treatment of Established Infections.

Animal	Unloaded POROPLASTIC®	Animal	Gentamicin POROPLASTIC®	Gentamicin Content (wt%)
71	6.2×10^7	61	0	6.2
72	2.0×10^7	62	0	6.2
73	3.1×10^7	63	0	6.2
74	ND	64	0	6.2
75	7.5×10^6	65	0	7.7†
76	5.0×10^7	66	0	7.8†
77	3.2×10^7	67	0	8.2†
78	2.9×10^7	68	0	8.4†
79	8.6×10^8	69	0	9.2†
80	3.8×10^8	70	0	9.4†

* 2.3×10^9 <u>S. aureus</u> Inoculated on Day One;
3.5×10^9 <u>S. aureus</u> Inoculated on Day Three.

†Gentamicin-POROPLASTIC® (Eight Months of Age).

ND: Not Done (Anesthesia Death).

Table 5. Concentration Of Gentamicin Sulfate In Wash Solutions From Wounds
Covered With POROPLASTIC®.

Group*	Type Of Dressing	Range Of Gentamicin Sulfate $(\mu g/cm^2)$	Wounds With Detectable Drug, %
I.			
a. First Dressing	Gentamicin-POROPLASTIC®†	2.07-30.69	100
b. Second Dressing	No Dressing	0.00--6.88	80
II.			
a. First Dressing	Gentamicin-POROPLASTIC®†	1.29-31.85	100
b. Second Dressing	Unloaded-POROPLASTIC®	0.00--2.75	40
III.			
a. First Dressing	Unloaded-POROPLASTIC®†	0.00	0
b. Second Dressing	Gentamicin-POROPLASTIC®	7.05-84.55	100

*First Dressing On Wound Three Days, Second Dressing On Wound An Additional Three
Days (n=10, For All Groups)

†6.2% w/w Gentamicin Sulfate

Serum removed from seven randomly selected guinea pigs (each treated three days previously with gentamicin-POROPLASTIC®) was also assayed for gentamicin sulfate. Gentamicin was detected in the serum of one animal at a concentration of 8.03 µg/ml. Very small zones of inhibition around discs impregnated with sera from two other animals indicated the presence of small quantities of gentamicin (2.45 and 1.87 µg/ml).

Results of the tissue culture test using Vero cells showed unloaded POROPLASTIC® to be noncytotoxic. At 96 hours post-exposure, a healthy monolayer of cells was observed bordering the POROPLASTIC®. Cytotoxicity was observed in wells containing POROPLASTIC® loaded with gentamicin; most probably due to the known cytotoxic effect of high concentrations of gentamicin sulfate on tissue culture cells.

DISCUSSION AND CONCLUSIONS

POROPLASTIC® is a versatile film that can be bonded to a wide variety of fibers and fabrics or can be formed as a coating on plastic films. Its physical characteristics indicate potential as an excellent wound dressing. In this study POROPLASTIC® has served as a successful vehicle for the sustained release of an antibiotic into infected wounds thereby resulting in excellent control of infection.

From work completed at this time, it has been determined that:

1. Gentamicin-POROPLASTIC® will effectively eliminate infections with the common wound infecting organism Staphylococcus aureus.

2. The rate of release of gentamicin from POROPLASTIC® film in vitro is dependent on loading and approximates zero order kinetics.

3. No adverse effects on gentamicin release kinetics were detected due to sterilization of POROPLASTIC® by gamma irradiation.

4. When covered with POROPLASTIC® and polyethylene, wounds incurred in the animal model remain hydrated with no eschar formation.

5. Gentamicin-POROPLASTIC® retains antimicrobial efficacy for at least eight months when stored at room temperature.

6. It is envisioned that a dual drug POROPLASTIC® system should

provide a wound dressing with broad spectrum efficacy.

Efforts are directed toward finding the best method for applying POROPLASTIC® to wounds. A viable alternative may be to position a commercially available polyurthane dressing such as Tegaderm® or Op-Site® over the POROPLASTIC®. Tegaderm® and Op-Site® will not adhere to guinea pig skin, therefore, clear polethylene (1600 mm², 5 mil) was positioned over the POROPLASTIC® and glued to surrounding intact skin with Histo-acryl®tissue adhesive (butyl-2-cyanoacrylate). This technique did not give sufficient adherence to prevent dressing loss for the three-day period. Use of Wonder Bond®, however, assured tight adherence and was used successfully. In future experiments, we envision using Wonder Bond® adhesive to attach the polyurethane dressing. Unlike polyethylene, poly-urethane should not negate the permeability characteristics of the POROPLASTIC®. The moisture vapor transmission rate, oxygen permeability characteristics, and known adhesiveness to human skin make Tegaderm® and OpSite® attractive as an outer dressing for POROPLASTIC® application to humans.

Serum gentamicin was detected in three of the seven animals bandaged for three days with gentamicin-POROPLASTIC®. Of these, verification of the two lower values is necessary due to the lack of definition and small size of inhibition zones resulting in the disc diffusion assay used. A serum gentamicin level of 8.03 µg/ml was detected in one animal. Serum gentamicin levels are affected by the dose administered and the drug excretion rate (kidney function). Since no relationship between the gentamicin serum levels and the presence of gentamicin-POROPLASTIC® with varying core-loads over the wounds of animals has been established, additional studies are being performed. Ideally, the wound dressing should control localized infections while producing minimal systemic antibiotic levels. The impressive success achieved in the total elimination of S. aureus with gentamicin-POROPLASTIC® (6.2 wt%) indicates the likely efficacy of dressings with lower drug concentations. Theoreti-cally, lower drug core-loads, when applied topically, should yield lower serum levels of drug.

It is of interest that wounds (Table 3) overlaid with gentamicin-POROPLASTIC® (6.2 wt%) for three days have, at the end of that time, similar amounts of gentamicin sulfate present on the wound surface (Table 5, Groups Ia and IIa). However, as shown in Table 5, wounds kept moist by rebandaging at three days with unloaded-POROPLASTIC® (Group II, b) yield wounds three days later with lower quantities of surface gentamicin sulfate (0.0-2.75 µg/cm², 60% drug-free) than wounds left air exposed (0.0-6.88, µg/cm², 20% drug-free). In light of the clinical observation that POROPLASTIC® (with polyethylene cover) keeps wounds hydrated and precludes echar formation, it is speculated that fluid exchange between

SETTERSTROM, VINCENT, and NICHOLS

the circulation and the POROPLASTIC® covered wound surface enhances the
entrance of gentamicin into the circulation for subsequent excretion.
Due to the stability of gentamicin sulfate, it is speculated that
trapping of drug within echars on the air-exposed wounds is enhancing
detection.

In conducting the research described in this report, the investiga-
tors adhered to the "Guide for the Care and Use of Laboratory Animals" as
prepared by the Committee on Care and Use of Laboratory Animals of The
Institute of Laboratory Animal Resources, National Research Council.

ACKNOWLEDGMENTS

REFERENCES

1. Hunt, T. K. and W. Van Winkle, Jr., 1979. Normal Repair, p. 42. In
T. K. Hunt and J. E. Dunphy (ed.), Fundamentals of Wound Management.
Appleton-Century-Crofts, New York.

2. Nichols, L. D. "Process of Preparing Gelled Cellulose Triacetate
Products and Products Produced Thereby." 1974. U.S. Patent 3,846,404.

3. Shriner, R. L., R. C. Fuson, and D. T. Curtin. 1959. The Systematic
Identification of Organic Compounds, 4th ed., p. 125. John Wiley & Sons,
New York.

4. Lantz, C. H., D. J. Lawrie, F. G. Witebsky, and J. D. MacLowry.
1980. Evaluation of Serum Gentamicin Assay Procedures for a Clinical
Microbiology Laboratory. J Clin Microbiol 12:583-589.

Microclimate Controlled (MC) Clothing for Extended Tank Crewman
Mission Time in CB/Hot Environments (U)

*THOMAS H. TASSINARI, MR. AND VINCENT D. IACONO, MR.
US Army Natick R&D Center
Natick, MA 01760

INTRODUCTION

Statement of the Problem

Crews of all types of armored vehicles and aircraft, operating in
extreme hot environments experience unusual problems of discomfort during
limited duration of exposure. Under adverse climatic and enemy imposed
conditions, contemporary combat protective clothing systems, even when in
combination with the ventilating or cooling device of the vehicle, fail to
provide conditions which will keep the crewman's body temperature within
reasonable limits. The crewmen receive very little benefit from cabin
ventilation in this situation because their bodies are isolated by their
combat clothing from the low flow of forced ventilating air.

To protect the Combat Vehicle Crewman (CVC) and ground soldier from a
multitude of enemy-imposed hazards occurring on the battlefield, effective
clothing systems for ground soldiers and combat vehicle crewmen have been
developed. However, as a result of these added protective layers of
clothing, a significant amount of environmental protection has been traded
off. The additional weight, increase in insulative properties, and
decrease in water vapor permeability of these clothing systems has
increased heat stress, especially during periods of activity in hot
environments.

The problem of heat stress is even more acute when the combat
vehicles are buttoned up to protect against chemical warfare (CW) agents.
Even with collective protection, the quantity of ventilating air in the
cabin will be greatly reduced. In all probability, the addition of
sweat from the crew will raise the humidity of the hot cabin air to
intolerable levels. Consequently, under these environmental conditions,

the crewmen will be subjected to severe heat stress causing a decay in their performance output. These deleterious physiological effects of high temperatures exert a powerful influence on the crewmen's abilities to operate their high performance equipment effectively. In the best case, a loss of mental performance will occur (1, 2) while in the worst case, heat stroke would be present, possibly leading to death. The end result at any rate is equipment malfunctioning and ultimately aborting the combat mission.

Impact of Hot Environments on Crewmen

Bodily heat regulation is essentially the maintaining of a balance between gains of heat on the one hand and losses on the other. The body is a machine, producing heat in proportion to its activity and method of operation, and is regulated by a control system which is complex and sensitive. The clothing and surrounding environment alter the body's heat balance, leading to heat stress in the individual.

Previous studies (3, 4, 5, 6, 7) demonstrated the magnitude of the heat stress problem in military personnel wearing Chemical Biological (CB) protective clothing and working in hot environments. For example, in a static desert test at Yuma Proving Ground, well trained and acclimated crewmen in CB protective clothing were able to function for only 80 minutes while buttoned up in an M1 tank. This test was performed on a relatively cool desert day with an internal tank temperature of $37.8^{\circ}C$ ($100^{\circ}F$). Depending on the environmental temperature/humidity, workload and level of CB and other types of protection, tolerance times can be expected to range from 15 minutes to 80 minutes in hot environments before a soldier becomes a heat-stress casualty.

Heat stress in the individual can lead to heat exhaustion or heat stroke maladies. Circulatory deficiency heat exhaustion is characterized by fainting or near collapse during or immediately following heat exposure while the subject is standing or working.

However, heat stroke is much more severe than heat exhaustion. It results from a failure of the body's heat loss mechanism to control core temperature. Physiological effects of the high fever produced by the uncontrollable rise in body temperature and heart rate include neurological and mental disturbances. Heat stroke can result in loss of consciousness, convulsion, delirium, and death.

Microclimate Cooling - A Viable Solution to the Heat Stress Problem

To extend the mission time of combat vehicle crewmen operating in CB threat environments, an active method of cooling is required. The two

244

alternatives are macroclimate cooling and microclimate cooling.
Macroclimate cooling is the conditioning of the living space while
microclimate cooling is the conditioning of the micro-environment between
the body surface and the inner clothing layer.

Macroclimate cooling is not practical for cooling of crewmen because
the size, weight, and power of the required refrigeration equipment are
greater than can be accommodated on board ground/air combat vehicles.
Another factor making this cooling method unacceptable is that the crewmen
are insulated against the macro-cooled environment by their protective
clothing. In comparison, microclimate cooling equipment is smaller, weighs
less, uses substantially less power, and provides efficient cooling to the
body. Furthermore, the insulation provided by the protective clothing
becomes an advantage by isolating the cool microenvironment inside the
clothing from the hot ambient conditions.

Results from Army funded studies (8, 9) of macrocooling indicate that
a four-man crew in a ground combat vehicle (for example, a tank) would
require 3 to 4 tons (10,548 to 14,064 W) of crew compartment cooling to
achieve only marginally acceptable vehicle cabin conditions in hot/dry
environments. In comparison, only 0.5 tons (1,758 W) would be required to
cool the same crew using a microclimate system.

Microclimate cooling system (MCS) configurations consist of three
types: ventilating air, conditioned air and chilled liquid. Each system
contains a cooling distribution garment, quick disconnects, umbilicals and
a source of cooling either a backpack or vehicle mounted console. The
garments are being designed to be of a minimum number of sizes and to be
compatible with standard combat clothing systems.

Of these three systems, chilled liquid and conditioned air are
effective in preventing heat stress in crewmen operating in high
temperature environments. With a ventilating air system, body cooling is
limited. Results from a recent study indicated that ventilating air was
ineffective under desert conditions but performed well for crewmen
operating in a tropic environment while working at a low rate (10).

Conditioned air and ventilating air systems are open loop, that is,
the air is discharged to the surrounding environment after heat exchange
with the body. On the other hand, the liquid system is closed loop which
means the coolant is continually recirculated between the cooling source
and the cooling garment.

The comparative characteristics of the two recommended microclimate
cooling types presented in Table 1 reveal their advantages and
disadvantages.

Program Approach

Current Army doctrine requires the protection of combat vehicle crewmen from nuclear, biological and chemical (NBC) agents. To meet this requirement, microclimate cooling systems will be incorporated into combat vehicles. Due to the large number of combat vehicles (ground and air), TRADOC HQ prepared and issued a prioritized list of vehicles requiring NBC/MCS capabilities. This list contains 32 ground and 8 air vehicles. Among these vehicles, there are a number of differences in chassis configuration, power train requirements, space allocation for new sub-components, crew complement and mission profile. All of these differences impact on the type/size of MCS selected for integration into the vehicle.

To help the man/vehicle/MCS integration, a front end MCS/vehicle trade-off analysis was conducted. The purpose of this investigation was to perform a systems analysis and hardware integration study to evaluate microclimate cooling (liquid/air) for use with a selected number of Army vehicles, each representing a specific category. The data developed from this study will be used for planning, selection, and procurement of the appropriate MCS for integration into combat vehicles. Results of this study identified air cooling as the preferred cooling system for most Army Combat Vehicles, with the exception of those vehicles which have a crew dismount requirement. While the selection of the air MCS is not optimum in terms of volume and power constraints, the importance of system reliability, sweat evaporation, comfort and compatibility with face mask integration makes air the preferred choice for most vehicles. On the other hand, the liquid MCS is the preferred system for the vehicles with a dismount requirement and the need to interface with portable liquid cooling backpacks.

In support to the M1E1 Product Improvement Program (PIP), an air-conditioned vest/connector was developed and tested. These items will serve as basic configurations for all other combat vehicles requiring an air conditioned MCS. This paper focuses primarily on laboratory design verification and field testing performed on this equipment.

MICROCLIMATE COOLING EQUIPMENT

Natick R&D Center is responsible for the development of microclimate clothing systems and equipment to meet the needs of all the Army's fighting men, especially the air/ground combat vehicle crewmen and infantrymen (11). To satisfy a diverse range of requirements, cooling garments, connectors, and backpack prototypes have been designed and built and are currently in various stages of development. Following is a brief description of this equipment including the M1E1 air cooled vest/connector.

Liquid Cooled Vest/Connector

The liquid cooled vest consists of a heat-sealed, polyurethane-coated nylon fabric bladder divided into four interconnected sections by stretch panels (for proper sizing and control). Flow channels direct a continuous stream of coolant through the bladder as heat is conducted away from the torso and neck. The vest weighs approximately 1.6 kg (3.5 lb) when full. The propylene glycol-water coolant is circulated through the vest at 22.7 kg/hr (50 lb/hr) and at temperatures of 10-16°C (50-61°F). A connector is being developed which enables the crewman to control the amount of cooling at the vest by adjustment of a liquid flow control valve. The connector also has a quick release coupling that seals automatically when disconnected to prevent coolant leakage. In addition, the connector is small, lightweight, and durable, and it will be low cost in production.

Liquid Cooling Backpack

The prototype liquid cooling backpack has a low profile with most of its weight supported on the hips. An ice cartridge is in the upper section of the pack, and a battery, reservoir, pump, and flow control hardware are in the lower section. This pack is capable of providing up to 2 hrs of continuous cooling to the crewman for dismount operations. The pack weighs 8.6 kg (20 lb) and supplies 45.4 kg/hr (100 lb/hr) of coolant at 10-13°C to the liquid cooled vest. When the ice is depleted, the specially designed ice cartridge is unplugged from the pack and plugged into a holding rack. Here it is refrozen in one hour by the rapid freeze refrigeration unit which can be located on the vehicle.

Other backpack cooling concepts are also being investigated. Consideration is being given to the use of various power sources such as batteries, gasoline engines and Stirling engines to drive a freon refrigeration unit with the entire system self-contained. Anticipated gains over the ice pack concept include (a) increased time of operation (up to 12 hours) between resupply of fuel for the prime movers and (b) independence from any vehicle mounted or otherwise located rapid freeze unit. Cooling of combat vehicle crewmen is simplified because the onboard

refrigeration unit would be smaller and lighter in weight and use less power if it were designed to cool the men while on the vehicle only. There would be no need for ice if one of these selfcontained concepts can be developed.

Air Cooled Vest/Connector for the M1E1 Abrams Tank

The M1E1 microclimate vest is designed to provide chest, neck, and back cooling via a hose/manifold system mounted on an open weave fabric (Figure 1). The hoses are lightweight and crush resistant and will maintain a constant inside diameter upon bending. Cooled air from the cooling unit is distributed by the connector so that 7.1×10^{-3} m^3/sec (15 scfm) is delivered to the vest and distributed at a ratio of approximately 40% to the chest, 20% to the neck and 40% to the back. The vest is lightweight (approximately 0.45 kg) (0.99 lb) and offers low resistance to airflow. It is worn over the undershirt and beneath the fragmentation protective vest (Figure 2).

Vest design allows the wearer to connect to the vehicle umbilical from the left or right side using the single air supply hose located in the center of the garment. This feature provides for maximum interface compatibility with the vehicle NBC system. Another design advantage is the bib configuration, i.e. the vest can be easily integrated with the fragmentation protective vest.

Figure 1   Natick Prototype Microclimate Cooling Air Vest for the M1E1

Figure 2   Crewman Wearing Natick Prototype Microclimate Air Cooled Vest Beneath the Fragmentation Protective Vest

The prototype air connector (Figure 3) is a multi-functional cooling component which interfaces with the primary and backup NBC protective systems. It splits 8.5×10^{-3} m^3/sec (18 scfm) of cooled air into 1.4×10^{-3} m^3/sec (3 scfm) for the ventilated facepiece and 7.1×10^{-3} m^3/sec (15 scfm) for the vest. Temperature control of vest cooling air is provided by a 7-position by-pass valve in the connector that restricts airflow to the vest while releasing the excess into the crew compartment. Other design features include low pressure drop through the connector, quick release when

Figure 3   Natick Prototype Microclimate Cooling Air Connector for the M1E1

disconnecting for rapid vehicle egress, and provisions for using the ventilated facepiece without the vest when the M13A1 backup NBC protective system is being utilized. The vest side of the connector is self sealing when disconnected to prevent the entry of chemical agents. Furthermore, the connector is fabricated of a lightweight, high strength engineering thermoplastic which can be injection molded for maximum production cost effectiveness.

M1E1 MICROCLIMATE COOLING RESULTS

A comprehensive test program was established to validate the M1E1 microclimate cooling system's readiness for deployment. First an engineering design test (EDT) of the vest and connector was conducted in August, 1983, at the Natick R&D Center Climatic Chambers. Then the entire system was tested during a dynamic field test at Yuma Proving Grounds (August - September 1983) as part of the M1E1 DT II.

In both the EDT and DT II, test subjects were dressed in the chemical protective CVC/MCS clothing ensemble (Figure 4) worn in the following order: underwear, fragmentation protective vest, CVC coverall, and chemical protective overgarment. To complete the combined clothing system, the M25A gas mask, chemical protective (CP) hood, CVC helmet, CP butyl boot cover and CP butyl gloves with cotton liners were also worn. The above chemical protective clothing items are worn when the individual is in mission oriented protective posture - 4 (MOPP-4), the highest level of chemical protection.

To insure safety of the test subjects by avoidance of heat stress disorders, physiological criteria established by the Surgeon General were used as a basis for terminating any test whether in the climatic chambers or out in the field. During any heat exposure, a rectal temperature above 39.5°C (103.1°F) and/or heart rate of 160 beats/min during rest and 180 beats/min during exercise for more than 5 minutes would lead to a halt of the test. Furthermore, subjective complaints such as headache, nausea, and dizziness would be reason to abort a particular test even if the heart rate and rectal temperature were within acceptable limits.

249

Engineering Design Verification Tests

A two-pronged study was conducted in the climatic chambers to verify the cooling effectiveness of the air-cooled vest prior to the start of the desert DT II for the M1E1. The first phase consisted of a 12-hour heat exposure at 49°C (120°F)/20% RH with the four tank crewmen generating heat at an average 240 W (819 Btu/hr) metabolic rate. This was accomplished by using a 50-minute walk/rest sequence on a level treadmill over the 12-hour exposure. For the second phase of testing, the crewmen generated an average 340 W (1160 Btu/hr) of metabolic heat over 3 hours by walking 3.6 Km/hr (2-1/4) mph on a level treadmill for 50 minutes (same as in 12 hour test) but only resting for 10 minutes each hour. In addition, a simulated solar load of 71 W (242 Btu/hr) was added using heat lamps hung from the ceiling.

Work rates used were based on what can be expected of a loader because he is the one with the most work in the tank. It is reasonable to conclude that if the cooling system meets the needs of the loader, then the other crewmembers will also be cooled adequately.

Ambient conditions of 49°C (120°F)/20% RH were selected because they are likely to occur inside a buttoned up M1E1 fighting in the desert on a hot day. Temperatures as high as 60°C (140°F) have been recorded inside vehicles without cooling systems, while parked in the desert sun.

For these tests conditioned air at 17°C (63°F) and 40% RH was supplied to each crewman. Air flow was 8.5×10^{-3} m^3/sec (18 scfm) to each crew station. The wye connector split off 1.65×10^{-3} m^3/sec (3.5 scfm) for the ventilated facepiece (VFP) while the remaining 6.8×10^{-3} m^3/sec (14.5 scfm) of air was supplied to the vest. The VFP must be supplied 1.4 to 2.1×10^{-3} m^3/sec (3.0 to 4.5 scfm) at all times in accordance with the NBC system requirements.

Physiological results indicate that crewmen dressed in MOPP-4 can survive for a least 12 hours under the severe conditions of this test. Average rectal temperature had increased only 0.8°C (1.4°F) during the test (Figure 5a) indicating a small amount of heat storage. An average exercise heart rate of 140 beats/minute at the test conclusion was indicative of minimal cardiovascular strain. During this test period, the subjects walked over 22.5 km (14 miles).

For the three hour stress test, average rectal temperature (Figure 5b) and heart rate were higher than for the 12-hour test. This situation reflects the increased heat stress resulting from the additional workload and supplementary thermal load (solar lights). All subjects completed the 3-hour test successfully with no apparent physiological problems. Final rectal temperatures ranged from 37.8°C to 39.2°C (100.0°F to 102.5°F) with
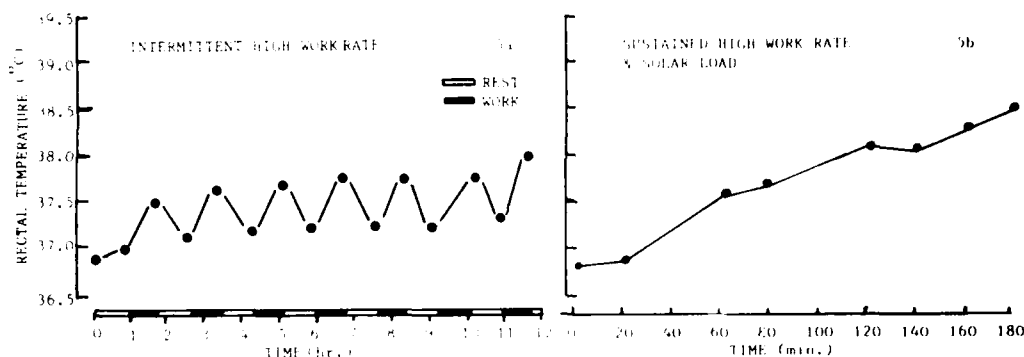
FIGURE 5. Average Rectal Temperatures of M1 Tank Crewmen During 12 hr. (5a) and 3 hr. (5b) Engineering Design Tests of MTDS.

an average of 38.5°C (100.4°F) while exercise heart rates ranged from 128 to 163 beats/min with an average of 150 beats/min at the end of the test.

Furthermore, each man performed an amount of work equivalent to loading 240 main gun rounds for the 12-hour test and worked hard enough to load 120 main gun rounds for the more stressful 3-hour test. Loading rates would be 20 rounds per hour for the 12 hour test and 40 rounds per hour for the 3-hour test. Without microclimate cooling, the predicted tolerance times are 80 minutes for the 3-hour test and 110 minutes for the 12-hour test (12). In other words, these tank crewmen dressed in MOPP4 and fighting in combat under similar environmental heat and workloads would become battlefield casualties if allowed to continue beyond the tolerance times. Microclimate cooling extended their mission time by more than seven-fold for the 12-hour test.

Subjective cooling responses of the crewmen during the two EDT phases are summarized in Table 2. A nine point rating scale ranging from 0 for unbearably cold to 8 for unbearably hot was used. In both tests, the torso was rated as the coolest body area and the hands as the most uncomfortable.

TABLE 2. Subjective Cooling Responses of M1 Tank Crewmen During Engineering Design Tests of the M1E1 Microclimate Cooling Vest Connector

BODY AREA	RATING	
	12 HR TEST	3 HR TEST
TORSO	COOL	--
CHEST	--	SL. COOL
BACK	--	SL. COOL
FACE	COMFORTABLE	COMFORTABLE
ARMS	WARM	WARM
HEAD	WARM	SL. WARM
LEGS	WARM	WARM
FEET	MOD. WARM	WARM
HANDS	HOT	VERY HOT
OVERALL	COMFORTABLE	COMFORTABLE

The overall rating was comfortable for both heat exposures even though the only area that was perceived to be cool was the torso. The

251

vest distributes conditioned air to this region first. Not much of
the air can travel to the extremities because it exits the uniform at
points of least resistance, such as the waist and neck.

Desert Development Tests

Two dynamic field tests were conducted at TECOM's Yuma Proving Ground
Facility. In the first test, a 12-hour mission was planned for the M1E1
but had to be aborted after 7.5 hours due to problems with the NBC cooling
system. The second test was added as a make-up and ran a full 12 hours with
all NBC cooling components operating as expected.

As indicated by the flatness of
the average rectal temperature
profiles (Figure 6), there was
essentially no heat storage in the
crew during either test. Sufficient
cooling was provided, even though, for
the 7.5 hour test, cooling air
temperature variations paralleled the
rise in desert ambient air
temperature. In fact, the loader
experienced air temperatures at his
vest from $28.9°C$ ($84°F$) at the start
of the test to almost $41.6°C$ ($107°F$)
at test conclusion in the afternoon.
In contrast, a properly operating NBC
cooling unit during the 12-hour follow-on test provided the loader's vest
cooled air at temperatures ranging from $23.8°C$ to $30.5°C$ ($75°$ to $87°F$).
Air flows for both tests were 7.1×10^{-3} m^3/sec (15 scfm) to the vest and
1.4×10^{-3} m^3/sec (3 scfm) to the ventilated facepiece.

FIGURE 6.   Average Rectal Temperatures Responses of
M1 Tank Crewmen During the 7.5 hr. and
12 hr. Desert Development Test II

Crewmen sweat rate data (Table 3) reveal that of the two heat
exposures, the 7.5 hour test was more heat stressful to the crew.   The

TABLE 3.   Sweat Rate Responses of Tank Crewmen During Desert Phase DT II at
Yuma Proving Ground of the M1E1 Microclimate Cooling System

TEST	SWEAT RATE $(G/M^2/HR)$		
	MAX	MIN	AVG
7.5 HR DESERT DT II	274	138	194
12.0 HR DESERT DT II	172	63	114

average sweat rate for the aborted test was 80 g/m^2/hr higher than for the full 12-hour test. More importantly, however, is that for either test the 18.9 L. (5 gallons) of drinking water carried inside the vehicle would be enough to sustain the crew. To replace the body water lost during the 12-hour test, 10.9 L. (2.89 gal) is required while if the 7.5-hour test continued to 12 hours, 18.6 L. (4.93 gal) would have to be replaced. In both trials, there was no dehydration evident in the crew as an average of less than 1% body weight was lost per man.

The crewmen had no major problems and stated they were comfortable during both dynamic tests. The only temporary discomfort was registered by the driver at one point in the 12-hour test when he was getting too much cooling. He was directed to change the setting on his wye connector to decrease the flow of air into the vest. This adjustment immediately corrected the problem, as was expected based on the connector design.

These results verify the operational capability and effectiveness of the M1E1 microclimate cooling system. Although the conditions of these field tests were not the severest in terms of workload and environmental temperatures, the engineering design test results previously obtained confirm the capability of this system to meet extreme heat stress challenges found in certain military critical regions of the world.

DISCUSSION

The results from the 7.5 hour desert test demonstrated that even though the conditioning unit failed to cool the air, crewmen performing light work were adequately cooled by sweat evaporation. This test suggests that ambient filtered air, circulated through the air cooled vest at less than 43.3°C (110°F), has the potential to serve as a back-up to the MCS and also to be used in support of vehicle silent watch. Ambient air cooling is currently being used successfully in the Army's Protective Outfit Toxicological Microclimate Controlled (POTMC) (13) for explosive ordnance disposal personnel.

During silent watch, a vehicle's engine and other systems are not operating and the crew is minimally active but in a state of readiness. Power for systems which are operating is being supplied by the batteries. It is critical during this period that sound from the vehicle be minimized and that battery power be conserved in order to restart the engine. Under this type of operational scenario, it appears that crewmen can be cooled by ambient air ventilation. This capability is a significant advantage for the conditioned air system. However, further work is required to determine if the available battery power on each vehicle can support the operation of the motor blower in the air conditioning unit during silent watch.

Due to the urgent need to provide individual cooling to the combat vehicle crewman, the Army (Natick R&D Center) investigated the suitability of off-the-shelf commercial microclimate cooling systems. Testing of these systems proved them to be inadequate from the standpoint of having insufficient cooling capacity, difficult to maintain, logistically unsupportable and low reliability (14). To correct the deficiencies would have required a major redesign effort without benefiting the Army's MCS program schedule. The Air Force has also reached similar conclusions with respect to these commercial systems.

In contrast to the critical deficiencies identified for the commercial systems, the microclimate cooling system described in this paper has a number of significant advantages. One of the most important attributes of this system is the capability to handle peak thermal loads (metabolic and ambient) as evidenced by the results of the successful 3-hour EDT. None of the commercial systems was able to pass this test.

Other advantages of this system include a one size-fits-all vest configuration (5th through 95th percentile), expected high reliability and long life as a result of using simple design and rugged materials, light weight and compatibility with standard CVC clothing items. Due to the anticipated low cost of the vest and connector in production, they will be throwaway items, thus eliminating the need for maintenance and spare parts. Using tape, the crewman can make temporary repairs to a leaking vest or connector until a replacement is obtained. From logistics and RAM-D viewpoints, these features are attractive and strongly support fielding of this system.

Impact on drinking water logistics is also favorable for the system. In a previous study (15), sweat rates of M1 tank crewmen were reduced by 50% when the microclimate air cooled vest was used as opposed to cooling the crew compartment. At the request of CG, TRADOC, the test was performed in an M1E1 tank under conditions simulating a hot day in Europe 32.2°C (90°F/74% RH). The results of this study are especially significant in desert/chemical warfare threat operations, where conservation of drinking water is imperative.

MOPP-4 encapsulation of 12 hrs has been adequately demonstrated. In light of these results, it appears possible that an extended MOPP-4 mission time for CVC can be accomplished. However, as we look to the soldier of the future and the trend to isolate him from his surrounding environment for long periods, a major problem area exists that must be considered; that is, the problem of duration of encapsulation. There is no apparent reason why the soldier of the future should not be able to maintain thermal balance with a portable MCS. The basic difficulty here will not be physiological but will be determined by the psychological

alterations of the soldier. The concern lies in the possibility that it will provide for an environment lacking in adequate sensory stimulation. A changing sensory environment appears to be essential for human beings. These concerns are under consideration and are an important part of our R&D program.

CONCLUSIONS

- Microclimate conditioned air (MCA) cooling is effective for at least 12 hours in preventing heat stress casualties of tank crewmen performing operational tasks in full chemical protective clothing in high temperature environments.

- MCA cooling is also effective in preventing heat stress casualties of tank crewmen performing high workload tasks over a 3-hour period.

- Under light workloads and environmental air temperatures of less than 43.3°C (110°F), clothing ventilation with ambient air by an MCA motor blower is capable of extending crew tolerance times. This has potential to support a vehicle silent watch mode of operations.

- The air-cooled vest and connector are designed for maximum compatibility with standard chemical protective/combat vehicle crewmen clothing systems and vehicle life support systems.

- Due to the favorable characteristics of air as a cooling medium, the following vest design features were achieveable: one size fits all (5th through 95th percentile); low physiological heat strain on the crewman when disconnected from the cooling unit; simple construction of low cost, lightweight materials; minimal maintenance required during long service life; and low cost, throw-away item. As a result, logistic support is simplified.

ACKNOWLEDGEMENT

The authors express their appreciation to the M1 Program Manager for the excellent support in providing funds and program coordination and to the M1 TRADOC Systems Manager at Ft. Knox for the superb cooperation received during this effort. In addition, the Military Ergonomics Division of the U.S. Army Research Institute of Environmental Medicine is recognized for the outstanding biomedical support provided during climatic chambers and field testing of the air cooled vest.

REFERENCES

1. Grether, W.F., "Human Performance at Elevated Environmental Temperatures," Aerospace Medicine, pp. 747-755 (July 1973)

2. Wilkinson, R.H.; Fox, R.H.; Goldsmith, R.; Hampton, E.F.; and Lewis, H.E.; "Physchological and Physiological Responses to Raised Body Temperatures," J. Appl. Physiol. 19:287-291 (1964)

3. Iacono, V.D.; Cohen, J.; Kupcinskas, M.; Gorbea, D.; and Fratantuono, J.; "Development of A Microclimate Cooling System," NLABS Tech. Rep. No. Natick/TR-82/010 (Feb 1982).

4. Henane, R.; Bittel, Jr.; Vinet, R. and Morino, S. "Thermal Strain Resulting from Protective Clothing of an Armored Crew in Warm Conditions," Aviat. Space Env. Med. 50:599-603 (1979)

5. Goldman, R.F. "Tolerance Time for Work in the Heat when Wearing CBR Protective Clothing," Mil. Med. 128:776-786 (1963)

6. Joy, R.T. and Goldman, R.F. "A Method of Relating Physiological and Military Performance: A Study of Some Effects of Vapor Barrier Clothing in a Hot Climate," Mil. Med. 133:458-470 (1968)

7. Spano, L.A. and Iacono, V.D. "Microclimate Controlled Protective Clothing Systems for Military Application," Proc. 1970 Army Science Conference, Ofc Ch. R&D, Dept of Army, Vol III, pp. 287-301 (1970), No. AD 713 581.

8. Wehner, J.M. "Combat Vehicle Cooling/Heating Design Investigation," Final Report - Contract No. DAAK70-80-C-0162, U.S. Army Mobility Equipment R&D Command, Sept 1981, No. AD-A115 109.

9. Dubay, M.C. "Engineering Analysis on Crew Compartment Heat Load and Auxiliary Power Unit Problems for the M-577/A1 Command Post Carrier with Collective Protection Equipment, " U.S. Army Tank Automotive Command, No. AD-B047851L, Mar 1970.

10. Iacono, V.D; Tassinari, T.H.; Kupcinskas, M.; Cohen, J.; Goff, R. and Fratantuono, J. "Performance Comparisons of Three Microclimate Systems for Cooling of Ground Combat Vehicle Crewmen in Chemical Protective clothing During Simulated Desert and Tropical Exposures," Natick R&D Center Tech Report, in press.

11. "Letter of Agreement for a Ground and Air Combat Vehicle Microclimate Conditioning System," USATRADOC ACN 61156, Jun 1981.

12. Givoni, B. and Goldman, R.F. "Predicting Rectal Temperature Response to Work Environment and Clothing" J. Appl. Physiol. 32:812-822, 1972.

13. McCullough, R.A., et al, "Final Report - Development Test II of the Chemical-Biological Protective Clothing Ensemble for Explosive Ordnance Disposal Personnel". Dugway Proving Ground AD No. B004910, Apr 1975.

14. Cosimini, H. et al, "A Determination of the Feasibility of Two Commercial Portable Microclimate Cooling Systems for Military Use" Natick R&D Center Tech Report, in press.

15. Toner, M.M. et al, "Comparison of Air Shower and Vest Auxiliary Cooling During Simulated Tank Operations in the Heat" USARIEM Tech Report No. T2/83.

A STABLE HIGH ACTIVITY PREPARATION OF DFPase FOR USE IN
DETECTION OF, AND PROTECTION AGAINST, NERVE AGENTS (U)

*CHARLES T. WHITE, Ph.D., CPT, MSC
ERIK J. MARTIN, Ph.D. AND KAREN A. UNRUH, MS.
LETTERMAN ARMY INSTITUTE OF RESEARCH
PRESIDIO OF SAN FRANCISCO, CALIFORNIA 94129

Chemical weapons may be used extensively in future wars. Recent press reports indicate that the Soviet Union and Soviet-backed countries have already employed chemical agents against Third World adversaries. Exposure of persons to mycotoxins, "yellow rain," which can be considered a chemical agent, has been reported in several regions of Indochina as well as in Afghanistan. Other news reports depict Soviet soldiers in full chemical protective clothing during military exercises in Europe indicating their capability to operate in a chemical environment.

Among the most feared chemical warfare agents are the nerve agents. Potential use of nerve agents poses threats to maintaining an effective fighting force. Full protective gear impairs performance in most of the soldier's normal duties. In addition, the threat of nerve agents can be a debilitating psychological weapon. Many of the Army's vehicles also become inoperable in a chemical environment due to inadequate filtering of outside air. Nerve agents not only impair the fighting force, but also have a deleterious impact on medical treatment facilities. If soldiers have been exposed to agents in the battle area, transportation and decontamination of patients must be performed with extreme care. Chemical casualties also drain medical resources since additional personnel are required for decontamination prior to medical treatment. Maintaining a clean treatment facility is another challenge. If nerve agents are used directly against medical treatment facilities, proper treatment of wounded soldiers would become virtually impossible.

There are still no effective antidotes for all nerve agents. Supportive therapy remains the best available treatment. Nerve agents inactivate cholinesterase. Cholinesterase is the enzyme which functions

at the neuromuscular junction and is responsible for regulation ot nerve conduction. Cholinesterase inhibited by the Soviet nerve agent, soman, "ages" rapidly and renders regeneration drugs, such as 2-PAM, ineffective. Soman also accumulates in body depots that are inaccessible to known antidotes. In contrast, the NATO nerve agent, sarin, undergoes aging slowly, thus 2-PAM is an effective antidote for sarin poisoning. The best defense against all nerve agents is rapid detection and adequate protection.

Current training manuals indicate the M-256 kit and the M-43/42 alarm system are the primary detection systems in the inventory. The M-256 kit is used at the squad level and can detect most known chemical agents. The kit consists of a placard containing numerous ampoules of chemicals and a supply of M-8 detecting paper. The placard requires numerous manipulations and takes 15 to 20 minutes to use; the M-8 paper requires only a few minutes. However, tests in this laboratory indicate that false positives may present a problem, e.g. DEET, the insect repellent currently used by the Army, was detected as a blister agent. DEET also caused the placard to indicate a positive for nerve agents. The M-43 detector with M-42 alarm are deployed with six per company and weigh 15 pounds each. In an Air-Land Battle 2000 scenario, 15-pound detectors would be unduly cumbersome.

Protection against nerve agents is currently provided by Mission Oriented Protective Posture (MOPP) suits. MOPP suits repel and/or absorb agents and thus must be decontaminated and/or changed every 6 hours. Personal decontamination is achieved by the M-258 kit which uses strong base and phenol, themselves severe skin irritants. Decontamination of equipment, such as tanks and other vehicles, is adequately performed with Super Tropical Bleach and caustics; however, sensitive equipment, i.e. electronics, may not tolerate these strong chemicals.

Enzymatic neutralization of nerve agents has potential use in both detection and detoxification. The only enzymes known to hydrolyze nerve agents are diisopropylfluorophosphatases (DFPases). Enzymatic detection has the added advantage of substrate specificity. Because the enzymes react with a limited number of substrates, false positives can be virtually eliminated. In the enzymatic hydrolysis of G-type nerve agents two moles of acid (H^+) and one mole of fluoride (F^-) are produced. This affords two possible variables to be monitored, changes in pH and fluoride concentration. Enzymatic hydrolysis is also rapid and quantifiable. Detoxification using enzymes should also be possible.

A solution containing an active enzyme can be applied or sprayed to decontaminate affected surfaces. Enzymatic hydrolysis is mild and works at biological pH's and temperatures. This may provide effective personal decontamination without irritation and the decontamination of sensitive equipment without damage. DFPase might also be packaged in barrier creams or attached to clothing to provide simultaneous protection and decontamination.

DFPase was first discovered by Mazur (1) at Edgewood Arsenal while studying the biological effects of the early nerve agents. This enzyme is found in the tissues of many species. Attempts were made to purify it. Instability of DFPase was the primary reason cited for failure to obtain pure high-activity preparations (2). In 1972 Hoskin and Long (3) reported the isolation and purification of a DFPase found in squid head ganglia. Although stable, only small amounts of this enzyme could be obtained. In 1975, the purification of porcine kidney DFPase (PK DFPase) was reported (4). The procedure was lengthy, time consuming, and resulted in a preparation of low specific activity. Instability also plagued this preparation of PK DFPase.

To assess the feasibility of enzymatic detection and decontamination a considerable and reliable supply of DFPase was required. Because porcine kidneys are available in fairly large quantity, we first set out to purify PK DFPase by taking advantage of substrate binding and charge differences, and second, to stabilize the enzyme after it was obtained pure.

Ion exchange chromatography was reported to destroy DFPase activity (2). However, by utilizing a DEAE bound to plastic beads, which affords high flow rates, ion exchange chromatography could be performed rapidly. This method reduced exposure time of the enzyme to the resin and afforded an enzyme preparation with activity much higher than that reported previously for any purified preparation of DFPase. This procedure was accomplished by using an initial DEAE column wherein the elution buffers were introduced in a stepwise fashion (Fig. 1a), followed by a second DEAE column eluted by a linear gradient (Fig. 1b). The enzyme preparation at this stage was still not pure.

The final purification employed affinity chromatography. A substrate for the enzyme, diisopropylphosphate, was covalently attached to an agarose matrix via an aminohexyl spacer arm. Aminohexyl agarose was treated with DFP in the presence of a proton scavenger dimethylaniline (DMA). This produced a column containing multiple
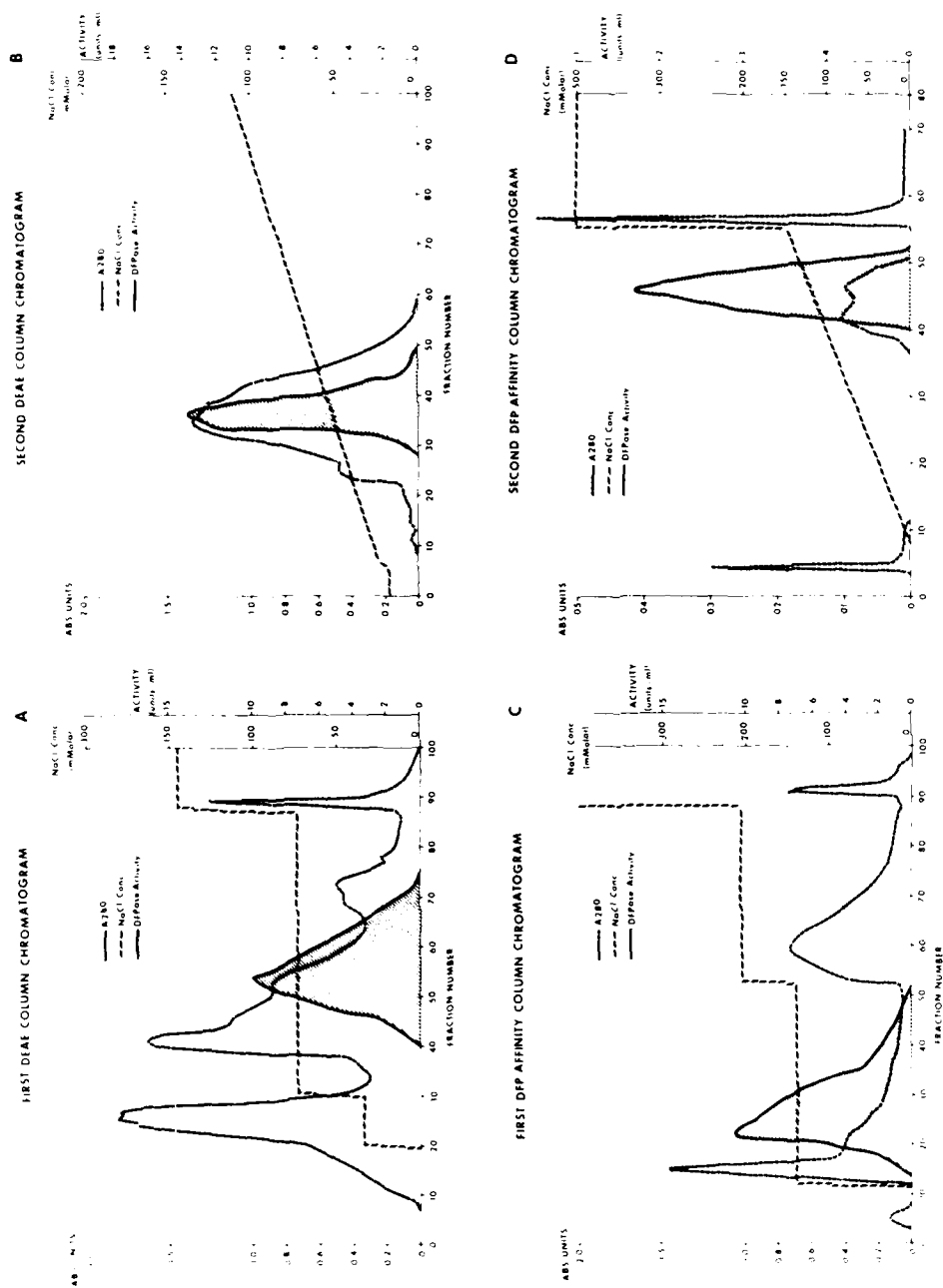
Fig. 1. Column Chromatograms. DEAE columns were 5.0 x 20 cm and eluted at a flow rate of 125 mL/hr. Substrate affinity columns were 2.5 x 15 cm and eluted at a flow rate of 80 mL/hr. Elution patterns were determined using a flow absorbance detector at 280 nm. In addition to the indicated NaCl concentrations, all elution buffers contained 25mM Tris, 1.0 mM MnCl$_2$, and 1.0 mM mercaptoethanol, pH 8.0. DFPase activity is indicated by the cross-hatched area.

262

substrate binding sites for the enzyme. Utilizing a two-step column procedure similar to that described for the ion exchange chromatography (Fig. 1c and 1d), we obtained a pure homogeneous enzyme with high specific activity. As shown in Table 1, a purification of over 4000-fold was achieved in just two columns with stepwise gradients

Table 1. PURIFICATION OF DFPASE FROM HOG KIDNEYS

FRACTION	Vol. ml	[1] Activity Units/ml	Protein mg/ml	[2] Sp. act. U/mg	Purification
CRUDE EXTRACT	1480	0.81	11.42	.071	—
60% FRACTION	830	1.03	3.23	.32	4.5
1st DEAE	30	76.44	9.38	8.15	114.8
2nd DEAE	13.2	199.92	7.33	27.3	384.5
1st DFP AFFINITY	9.3	286.7	2.05	139.9	1970
2nd DFP AFFINITY	6.3	264.6	0.75	352.8	4969

[1] 1 unit = 1 μ mole DFP hydrolyzed/minute.

[2] Activities measured at RT (22°C)

followed by linear gradient. The activity was determined with DFP as the substrate, however, analysis by other laboratories demonstrated that DFP and soman were hydrolyzed at the same rate. This DFPase preparation is twenty times more active than any so far reported (Fig. 2). The final product is a single band by sodium dodecyl sulfate acrylamide gel electrophoresis analysis with a molecular weight of 71,000.
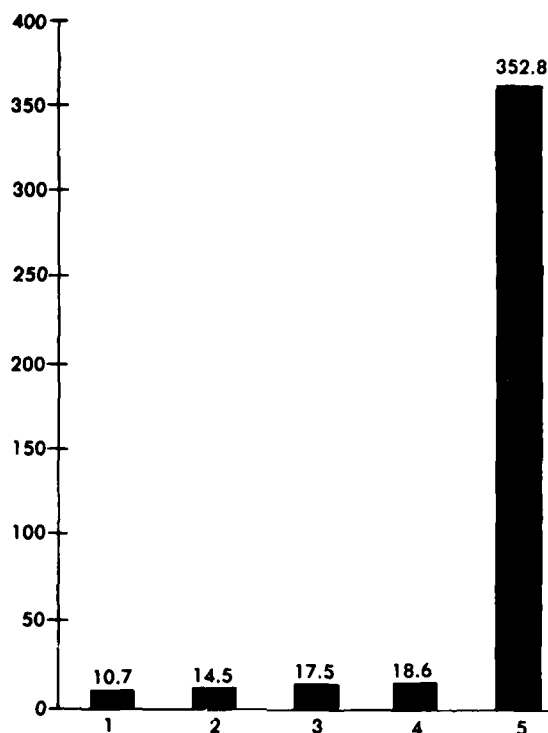
1. Storkebaum and Witzel. Forschungsberichte des Landes Nordrhein-Westfalen 1975;2523.
2. Mounter, Floyd and Chanutin. J Biol Chem 1953;204:221.
3. Cohen and Waringa. Biochim Biophys Acta 1957;26:29.
4. Hoskin and Long. Arch Biochim Biophys 1972;150:548.
5. White, Martin and Unruh 1983.

Fig. 2. Relative Specific Activities of Purified DFPases.

Despite the rapid purification with high degree of purity, instability of PK DFPase was still a problem. Storage of PK DFPase at 4°C tor 7 weeks resulted in a loss of 85 percent of the original activity. Numerous media and additives were evaluated as stabilizers of DFPase (Fig. 3). In a medium containing 25 percent or greater concentrations of glycerol, PK DFPase can be stored for 6 months at 4°C with greater than 80 percent retention of activity.
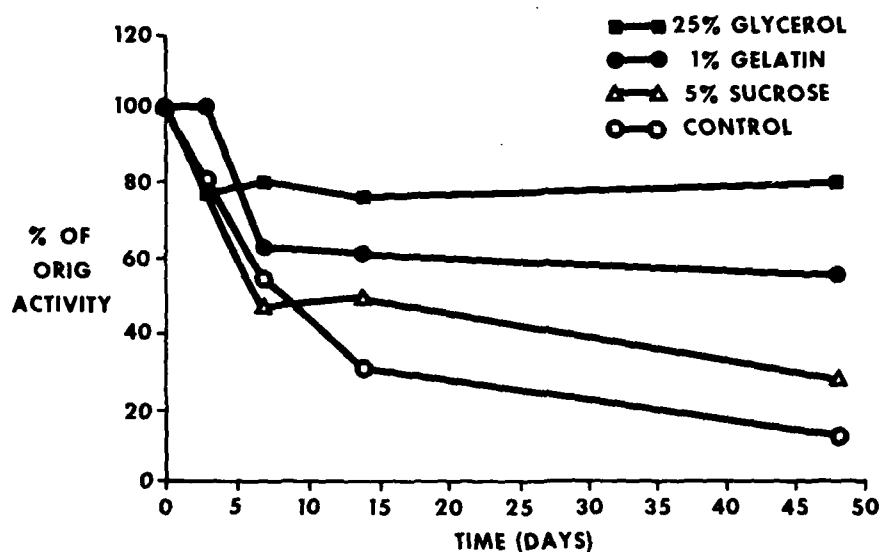
Fig. 3. Effect of Additives on DFPase Activity During Storage

Development of high-flow ion-exchange chromatography and substrate affinity chromatography for the purification of DFPase coupled with the discovery of a stabilizing medium has afforded, for the first time, a pure, reliable supply of highly active DFPase. Evaluation of the feasibility of enzymatic detection and detoxification is now possible. DFP was used in these studies because it is an analog for G-type nerve agents and, like nerve agents, contains a high energy P-F bond. Hydrolysis of this anhydride bond, as mentioned, generates two moles of acid. Phenol red, an acid base indicator, changes from red to yellow in a pH range where DFPase is active. Therefore, addition of DFP to a solution of DFPase and phenol red causes the solution to turn from red to yellow within 1 to 2 minutes. Under substrate-limited conditions the rate of appearance of yellow color would be proportional to the amount of active nerve agent allowing for the quantitation of nerve agent. This pH concept can also be adapted for a detection paper. A solution of the enzyme and phenol red was absorbed on filter paper. As shown in Figure 4 both liquid and vapor were detected.
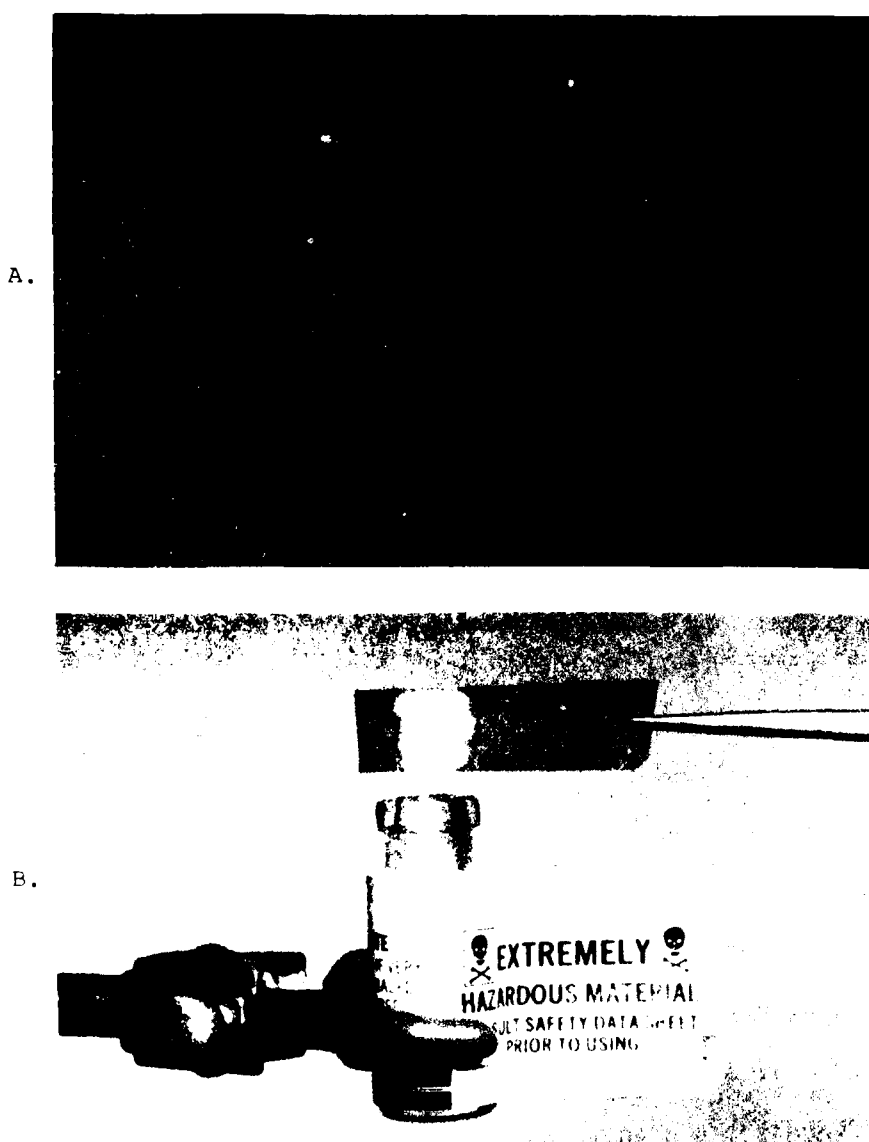
Fig. 4. a) Appearance of the test paper 2 minutes after adding 20 μl of a 10^{-4} M solution of DFP. Test paper was prepared by first saturating with a solution of phenol red in 25 mM Tris at pH 8.5 and allowed to air dry. The paper was then wetted with a 1 mg/mL solution of DFPase in 25 mM Tris pH 8.5. b) Appearance of the test paper 2 min after suspension over a vial of DFP opened in a chemical fume hood.

Electronic detectors that sense changes in fluoride concentrations are being investigated. Since DFP hydrolyzes G-type nerve agents, solutions of DFPase may be useful in decontaminating these nerve agents. It may also be possible to incorporate DFPase into barrier creams or clothing for protection.

This research has demonstrated that it is possible to obtain a pure, highly active, stable preparation of DFPase from hog kidneys, and that it is feasible to use this preparation to detect organophosphates in both the liquid and vapor phases. In order to carry out more elaborate studies focusing on the use of DFPase in barrier creams and bound to clothing, a cheap, even more stable source of DFPase must be found. A number of species of bacteria produce DFPase, and with the aid of genetic engineering, yields should be improved. If DFPase can be produced in industrial quantities, research can determine its practical uses in detection, protection, and decontamination of G-type nerve agents.

REFERENCES

1. Mazur A. An enzyme in animal tissues capable of hydrolyzing the phosphorus-fluorine bond of alkyl fluorophosphates. J Biol Chem 1946;164:271-289.
2. Cohen JA, Warringa MGPJ. Purification and properties of dialkylfluorophosphatase. Biochim Biophys Acta 1957;26:29-34.
3. Hoskin FCG, Long RJ. Purification of a DFP-hydrolyzing enzyme from squid head ganglion. Arch Biochem Biophys 1972;150:548-555.
4. Storkebaum W, Witzel H, Ruterjans H. Isolation, characterization, and analysis of a phosphoric acid triester hydrolyzing enzyme isolated from pig kidneys. Forschungsberichte des Landes Nordrhein-Westfalen 1975;2523:1-22.

ESTIMATING PRECISION AND BIAS AS COMPONENTS

OF ACCURACY IN A DYNAMIC MEASUREMENT SITUATION

NORMAN L. WYKOFF, PHD
US ARMY JEFFERSON PROVING GROUND
MADISON, INDIANA 47250

1. Introduction: The problem of determining the accuracy (nearness to the truth) of measurements and measurement systems is one that has plagued mankind in the past, but never to the extent that it does now or will in the future. Measurement errors that were not even noticed in the past are now considered to be intolerable. Unfortunately, accuracy is a very elusive concept. It is easy to discuss and in fact to understand, but to date, no reasonable quantification is known.

In order to understand accuracy, it is first necessary to discuss two more primitive concepts, namely bias and precision. Bias is the tendency to measure something other than the true value such as one might do with a cheap ruler or a bent micrometer or a high quality instrument when one fails to make temperature corrections. Bias is the difference between the observed value and the true value which is often unknown and unknowable.

The precision of a measurement process is related to reproducibility, the ability to yield virtually the same result when repeating a measurement. It is a measure of closeness together of the results upon repetition of the experiment. Imprecision is ordinarily summarized by the standard deviation of the process, i.e., a large standard deviation implies an imprecise system, the measurements are not "close together".

To be considered accurate, a system must not be biased, i.e., it must report the true value, and it must be precise, that is, all the values it reports must be close to the true value. The illustration is often given of the marksman who shoots a pattern of five shots in a one inch circle six inches from the bullseye. Clearly, his shooting is very precise. However, the bias error of six inches means that his shooting is not accurate. Meanwhile, another marksman sprays five shots over a two foot circle around the bullseye in such a way that the center of the pattern is the bullseye. In this case, the system is unbiased, but very imprecise and therefore again not accurate. Accuracy has to do with closeness to the truth, precision only with closeness together, and

bias the difference between the average observed value and the true value.

Three situations seem to arise in practice. In the first, some "absolute standard" such as a gage block is available. The equipment is exercised yielding a set of measurements. The difference between the average reading and the "true value" can be calculated since this characteristic is known for the gage block. This is the bias for the system. If the bias is not zero, typically one makes adjustments so that it will be zero, i.e., the system is "calibrated". The standard deviation is calculated for the set of measurements and is used as a measure of the imprecision of the system.

In the second situation, no absolute standard exists; the best one can do is a relative comparison of systems. Measurements are taken of the test items with each of the competing systems. Using certain simplifying assumptions, one can calculate the difference in the bias between systems and the imprecision of each system.

The third situation is extremely general; no absolute standard exists, in fact it is not even necessary to measure exactly the same events. In this context little if anything can be said about the bias. However, very good bounds can be calculated for the imprecision. This is the model for the major results of this paper.

The particular problem that triggered this study is the measurement of the velocity of projectiles in flight and the evaluation of available measurement instruments. Since this is a dynamic situation, repetition of measurements is not possible, there is only one chance to obtain a measurement. Further, it is a confounding characteristic of the equipment currently available that sensors cannot be placed coincidentally. This factor will be ignored in the analysis for the following reasons. First, the placement while not coincident may be nearly so, thus only rarely contributing any significant error. Second, the restriction may not be true for future equipment. Finally, velocities are usually extrapolated back to the muzzle of the weapon, thereby supposedly removing concern about the noncoincidence of the sensors. However, one more comment must be made regarding the correction back to the muzzle. This introduces its own errors into the reported velocity, but in those cases where the sensors are close to each other, these corrections are virtually identical and thus will affect the error terms in the same way.

2. <u>General Model</u>: Let T_{ij} be the true value, X_{ij} the observed

value, e_{ij} the random error (assumed to have mean zero and to be

independent from T_{ij} and E_{ij}), and E_{ij} the nonrandom error in the j-th

measurement with the i-th instrument, $i = 1, \ldots, k$ and $j = 1, \ldots, n$.

Then
$$X_{ij} = T_{ij} + E_{ij} + e_{ij} \qquad\qquad (2.1)$$

where the values of the individual terms in the RHS are unknown. The ultimate goal of this study is to estimate parameters of the distribution of error terms for each instrument and thereby make inferences about bias, precision and accuracy of the instrument. This paper is only concerned about making point estimates of bias and precision, the other topics will follow later.

A note of definition is in order regarding the terms bias, precision, and accuracy. A review of the literature yields a variety of notions regarding the terms precision and accuracy. The purpose of this study is not to resolve such arguments but is rather to obtain reasonable procedures for estimation of bias, precision, and accuracy as defined below.

<u>Definition</u>: The bias of the population of measurements for the i-th instrument characterized in (2.1) is the mean or expected value of the nonrandom errors. Notationally:

Bias of the i-th instrument $= \mu(E_i) = E(E_i)$

<u>Definition</u>: The precision of the population of measurements for the i-th instrument characterized in (2.1) is the inverse of the standard deviation of the random errors. Notationally:

Precision of the i-th instrument $= [\sigma(e_i)]^{-1}$.

Note that this definition implies that imprecision is directly proportional to $\sigma(e_i)$, i.e. if $\sigma(e_i)$ is large (small) so is the

imprecision. Consequently, given two instruments I_1 and I_2 with $\sigma(e_1)$

and $\sigma(e_2)$, we say that I_1 is more precise than I_2 if $\sigma(e_1) < \sigma(e_2)$.

<u>Definition</u>: An instrument or a set of measurements is accurate if it is unbiased and precise.

3. <u>Estimating Precision</u>: The three cases described in the introduction can now be described more explicitly.

<u>Case I</u>. $\qquad T_{ij} = C$, where C is a known constant
and
$$E_{ij} = E_i, \quad i = 1, \ldots, k; \ j = 1, \ldots, n.$$

now
$$\mu(X_i) = E(X_{ij}) = C + E_i$$
or
$$E_i = \mu(X_i) - C,$$

so the Bias, E_i is estimable.

Also for the variances,

$$\sigma^2(X_i) = \sigma^2(e_i),$$

so the imprecision is also estimable.

<u>Case II.</u> $\quad T_{ij} = T_{i'j}, \ i \neq i'=1, \ \ldots \ , k; \ j=1, \ \ldots \ ,n.$

Now

$$\mu(X_i - X_{i'}) = \mu(E_i - E_{i'})$$

and

$$\sigma^2(X_i - X_{i'}) = \sigma^2(E_i - E_{i'}) + \sigma^2(e_i - e_{i'}).$$

‾n this case differences in bias errors and bounds on precision are estimable. Further, the precision is estimable as will be shown later.

<u>Case III.</u>

Now

$$\mu(X_i - X_{i'}) = \mu(T_i + E_i - T_{i'} - E_{i'}) \tag{3.1}$$

and

$$\sigma^2(X_i - X_{i'}) = \sigma^2(T_i + E_i - T_{i'} - E_{i'}) + \sigma^2(e_i - e_{i'}). \tag{3.2}$$

Obviously, $\quad \sigma^2(e_i - e_{i'}) \leqslant \sigma^2(X_i - X_{i'}) \tag{3.3}$

with equality iff $\ T_{ij} + E_{ij} - T_{i'j} - E_{i'j}$ is constant for all $j = 1,$

$\ldots, n.$ It is also true that

$$\sigma^2(e_i + e_{i'}) = \sigma^2(X_i + X_{i'}) - \sigma^2(T_i + E_i + T_{i'} + E_{i'})$$

or that

$$\sigma^2(e_i + e_{i'}) \leqslant \sigma^2(X_i + X_{i'}) \tag{3.4}$$

but the RHS is usually considerably larger than the LHS since in most cases of interest, the true values are the same or nearly so.

When the random errors are independent of each other,

$$\sigma^2(e_i \pm e_{i'}) = \sigma^2(e_i) + \sigma^2(e_{i'})$$

so

$$\sigma^2(e_i) + \sigma^2(e_{i'}) \leqslant \sigma^2(X_i \pm X_{i'}) \tag{3.5}$$

with equality iff $T_{ij} + E_{ij} \pm T_{i'j} \pm E_{i'j}$ is constant, $j = 1, \ldots, n.$

The case of equality in (3.5) with $\sigma^2(X_i - X_{i'})$ in the RHS has been extensively covered by Grubbs (1948), Thompson (1963), Maloney and Rastogi (1970), Grubbs (1973) and Hanumara (1975).

Focusing on the first instrument, we have from (3.5),

$$\sigma^2(e_1) \leq \min \{\sigma^2(X_1 \pm X_{i'}), \ i' = 2, \ldots, k\}$$

and just as obviously,

$$\sigma^2(e_i) \leq \min \{\sigma^2(X_i \pm X_{i'}), \ i \neq i' = 1, \ldots, k\}, \ i = 1, \ldots, k \quad (3.6)$$

subject only to the restriction that the random errors are independent of each other and of the true and non-random error values. Since we obviously do not generally know $\sigma^2(X_i \pm X_{i'})$, we must base our estimate on the sample variance $s^2(x_i \pm x_{i'})$. But now the inequality may no longer hold. Therefore, we make the following approximation, which yields an estimate that is likely to be large for, at least, large samples. The estimate for $\sigma^2(e_i)$ when e_i and $e_{i'}$ are independent is:

$$\hat{\sigma}_1^2(e_i) = \min \{ s^2(x_i \pm x_{i'}), \ i \neq i' = 1, \ldots, k \}, \ i = 1, \ldots, k \quad (3.7)$$

Example: The following coded data is from an actual firing with 3 independent velocity measuring systems.

Rd No.	SYSTEM A	SYSTEM B	SYSTEM C
1	51.45	45.82	78.05
2	21.45	16.14	48.91
3	95.89	90.77	123.78
4	105.04	99.99	133.89
5	128.93	123.84	157.17
6	115.85	110.92	145.21
7	130.34	125.39	158.94
$\bar{x}$	92.70	87.55	120.85
s^2	1700.90	1716.81	1758.21

The differences in mean is explained by the fact that these are instrument velocities and that the systems are at different distances from the weapon. It should also be noted that systems A and B are electrical and essentially the same, while C is an optical device several hundred feet closer to the muzzle than A or B.

Forming the differences we obtain:

	A − B	A − C	B − C
1	5.63	−26.60	−32.23
2	5.31	−27.46	−32.77
3	5.12	−27.89	−33.01
4	5.05	−28.85	−33.90
5	5.09	−28.24	−33.33
6	4.93	−29.36	−34.29
7	4.95	−28.60	−33.55
$\bar{x}_2$	5.15	−28.15	−33.30
s^2	.06	.85	.49

Therefore, from (3.7)

$$\hat{\sigma}_1^2(e_A) = .06$$

$$\hat{\sigma}_1^2(e_B) = .06 \qquad\qquad (3.9)$$

$$\hat{\sigma}_1^2(e_C) = .49$$

which indicates that systems A and B may be superior to C in regard to precision, however, since in (3.8) we only have estimates, care must be exercised in making such assertions.

4. Special Models.

4.1 Constant Model: Suppose $T_{ij} + E_{ij} \pm T_{i'j} \pm E_{i'j}$ is constant

for all choices of i, i', and j, then (3.5) is

$$\sigma^2(e_i) + \sigma^2(e_{i'}) = \sigma^2(X_i \pm X_{i'}).$$

Further, as Grubbs (1948) showed, (call the following the Grubbs variance)

$$\sigma_C^2(e_i) = \frac{1}{k-1} \left[\sum_{i' \neq i} \sigma^2(X_i \pm X_{i'}) - \frac{1}{k-2} \sum_{i' \neq i \neq i''} \sigma^2(X_{i''} \pm X_{i'}) \right] ,$$

i=1,...,k, where the subscript C is used to denote the constant model. Grubbs estimate then is:

$$\hat{\sigma}_C^2(e_i) = \frac{1}{k-1} \left[\sum_{i' \neq i} s^2(x_i \pm x_{i'}) - \frac{1}{k-2} \sum_{i' \neq i \neq i''} s^2(x_{i''} \pm x_{i'}) \right] ,$$

i=1,...,k. (4.1.1)

In particular for k=3,

$$\hat{\sigma}_C^2(e_1) = 1/2 \ [s^2(x_1 \pm x_2) + s^2(x_1 \pm x_3) - s^2(x_2 \pm x_3) \ ] \text{ etc., } \quad (4.1.2)$$

and for k = 4,

$$\hat{\sigma}_C^2(e_1) = 1/3 \ \{ \ s^2(x_1 \pm x_2) + s^2(x_1 \pm x_3) + s^2(x_1 \pm x_4)$$

$$- 1/2 \ [s^2(x_2 \pm x_3) + s^2(x_2 \pm x_4) + s^2(x_3 \pm x_4) \ ] \ \} \text{ etc.} \quad (4.1.3)$$

These are very nice concise results, however, it will be shown they are very sensitive to departures from the model. A major problem with the use of Grubbs' estimators is that very often one obtains a negative

value for $\hat{\sigma}_C^2(e_i)$ for some i. Techniques have been suggested to

obviate such results, but they are not entirely satisfactory. This problem can be alleviated by using the more general model of this paper.

4.2 Proportional Model

Suppose $T_{ij} + E_{ij} = k_i(T_{.j} + E_{.j})$, $k_i > 0$ for all i and j. Then

$$\sigma^2(e_1) + \sigma^2(e_2) = \sigma^2(X_1 - X_2) - (k_1 - k_2)^2 \ \sigma^2(T + E) \text{ etc, say.}$$

Hence if k = 3,

$$\sigma^2(e_1) = \tfrac{1}{2} \ [\sigma^2(X_1 - X_2) + \sigma^2(X_1 - X_3) - \sigma^2(X_2 - X_3)] \quad (4.2.1)$$

$$-\tfrac{1}{2} \ \sigma^2(T+E) \ [(k_1 - k_2)^2 + (k_1 - k_3)^2 - (k_2 - k_3)^2] \quad \text{etc.}$$

Now suppose (without loss of generality since one can permute the subscripts) that $k_1 \leqslant k_2 \leqslant k_3$, then,

$$(k_1 - k_2)^2 + (k_1 - k_3)^2 - (k_2 - k_3)^2 \geqslant 0$$

$$(k_1 - k_2)^2 - (k_1 - k_3)^2 + (k_2 - k_3)^2 \leqslant 0$$

$$-(k_1 - k_2)^2 + (k_1 - k_3)^2 + (k_2 - k_3)^2 \geqslant 0.$$

Therefore,

$$\sigma^2(e_1) \leqslant \tfrac{1}{2} [\sigma^2(X_1 - X_2) + \sigma^2(X_1 - X_3) - \sigma^2(X_2 - X_3)]$$

$$\tfrac{1}{2} \ [\ \sigma^2(X_1 - X_2) - \sigma^2(X_1 - X_3) + \sigma^2(X_2 - X_3)] \leqslant \sigma^2(e_2) \quad (4.2.2)$$

$$\sigma^2(e_3) \leqslant \tfrac{1}{2} [- \sigma^2(X_1 - X_2) + \sigma^2(X_1 - X_3) + \sigma^2(X_2 - X_3)].$$

That is,
$$\sigma^2(e_1) \leqslant \sigma_C^2(e_1)$$
$$\sigma_C^2(e_2) \leqslant \sigma^2(e_2)$$
$$\sigma^2(e_3) \leqslant \sigma_C^2(e_3).$$

To reiterate, the Grubbs' variance is a bound for the imprecision for the proportional model, in particular, the Grubbs' variance is a lower bound for $\sigma^2(e_2)$.

In this model a negative value for the LHS of the second inequality above may not give any useful information, but it also causes no concern and in fact is very reasonable. Hence one can see that very slight deviations from Grubbs' model can lead to spurious results when using Grubbs' estimators.

As one might imagine, similar results hold when k > 3. It is also informative when considering k = 4 to look at the Grubbs' variances based on 3 of the 4 instruments. One finds that $\sigma^2(e_1)$ is bounded above by each of the Grubbs' variances based on 3 or 4 instruments with $k_1 \leqslant k_2 \leqslant k_3 \leqslant k_4$ as well as the bound given in (3.6). However, for $\sigma^2(e_2)$ we have

$$\tfrac{1}{2} [\, \sigma^2(X_1 - X_2) - \sigma^2(X_1 - X_3) + \sigma^2(X_2 - X_3)] \leqslant \sigma^2(e_2)$$
$$\tfrac{1}{2} [\, \sigma^2(X_1 - X_2) - \sigma^2(X_1 - X_4) + \sigma^2(X_2 - X_4)] \leqslant \sigma^2(e_2)$$
$$\sigma^2(e_2) \leqslant \tfrac{1}{2} [\, \sigma^2(X_2 - X_3) + \sigma^2(X_2 - X_4) - \sigma^2(X_3 - X_4) \,]$$
$$\sigma^2(e_2) \leqslant \min \, \{ \, \sigma^2(X_2 - X_i), \, i = 1, \, 3, \, \ldots, k \}.$$

Note that the Grubbs' variance based on all four instruments does not appear, it is some times smaller and sometimes larger than $\sigma^2(e_2)$ depending on the values of the k_i's. The situation for $\sigma^2(e_3)$ is like $\sigma^2(e_2)$ and $\sigma^2(e_4)$ is like $\sigma^2(e_1)$. Investigation for k > 4 yields similar results.

5. <u>Dependence Among Some e_i's</u>: In the previous material we have assumed that the e_i's were independent of each other and of the nonrandom values. We can relax that restriction very slightly as follows:

Suppose, for example, that e_1 and e_2 are not independent of each other. Then (3.3) is still true since no use was made of the assumption that e_1 and e_2 are independent of each other until later.

Therefore it is true that $\sigma^2(e_1 - e_2) \leq \sigma^2(X_1 - X_2)$.

Of course, it is still true that $\sigma^2(e_1) + \sigma^2(e_i) = \sigma^2(e_1 - e_i)$ if e_1 and e_i are independent, so that $\sigma^2(X_1 - X_i)$ can still be used as an upper bound for $\sigma^2(e_1)$ if $i \neq 2$. Also, since

$$COV(e_1, e_2) \leq \sigma(e_1) \; \sigma(e_2) \text{ and } (3.3) \text{ yield}$$

$$|\sigma(e_1) - \sigma(e_2)| \leq \sigma(e_1 - e_2) \leq \sigma(X_1 - X_2)$$

and therefore

$$\sigma(e_1) \leq \sigma(e_2) + |\sigma(e_1) - \sigma(e_2)| \leq \sigma(e_2) + \sigma(X_1 - X_2) \qquad (5.1)$$

which may be helpful if one has an independent estimate of $\sigma(e_2)$.

Hence, given at least one instrument that is independent of the others, one can still estimate the errors for each instrument.

6. <u>Shared Sensors</u>: Velocities are always measured over an interval, no current equipment can yield an instantaneous velocity. Typically, two sensors are placed a known distance apart along the trajectory and the time of flight over the interval is measured. The average velocity over the interval is then calculated from the familiar formula $D=rt$. Often, it is desirable to use more than two, say m, sensors and calculate the velocities determined by the $\binom{m}{2}$ pairs.

The goal now is to estimate the random error of each sensor rather than of the pairs. To that end, consider the random error in the system consisting of sensors i and j, call it e_{ij}. This random error is the sum of the random errors in each of the sensors, so let $e_{ij} = \xi_i + \xi_j$. Quite obviously, if ξ_i and ξ_j are independent of each other, $\sigma^2(e_{ij}) = \sigma^2(\xi_i) + \sigma^2(\xi_j)$, i, $j=1,\ldots,m$. $\qquad (6.1)$

END

FILMED

DTIC
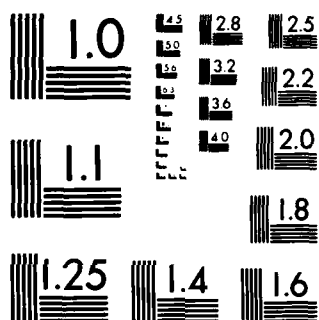
MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS 1963 A

As in earlier sections, we examine the variance of the difference of errors in the two systems, i.e. $\sigma^2(e_{ij} - e_{i'j'})$. However,

$$\sigma^2(e_{ij} - e_{i'j'}) = \sigma^2(\xi_i + \xi_j - \xi_{i'} - \xi_{j'}) \quad (6.2)$$

so that if $j = j'$, we have, for ξ_i and $\xi_{i'}$ independent,

$$\sigma^2(e_{ij} - e_{i'j}) = \sigma^2(\xi_i - \xi_{i'}) = \sigma^2(\xi_i) + \sigma^2(\xi_{i'}). \quad (6.3)$$

Of course, if i, i', j, j' are all different and independent,

$$\sigma^2(e_{ij} - e_{i'j'}) = \sigma^2(\xi_i) + \sigma^2(\xi_j) + \sigma^2(\xi_{i'}) + \sigma^2(\xi_{j'}). \quad (6.4)$$

In any case, analogous to the earlier discussion, for ξ_i independent of ξ_j, $i, j = 1, \ldots, m$,

$$\sigma^2(\xi_i) \leq \min\{ \sigma^2(X_{ij} \pm X_{i'j'}), \ i \neq i'; \ i \neq j \neq i'; \ i \neq j' \neq i'; \quad (6.5)$$

$i', j, j' = 1, \ldots, m\}$.

and we take

$$\hat{\sigma}_2^2(\xi_i) = \min\{ s^2(X_{ij} \pm X_{i'j'}), \ i \neq i'; \ i \neq j \neq i'; \ i \neq j' \neq i'; \quad (6.6)$$

$i', j, j' = 1, \ldots, m\}$.

Obviously, having calculated $\hat{\sigma}_2^2(\xi_1)$ and $\hat{\sigma}_2^2(\xi_2)$, say, one can add them to derive $\hat{\sigma}_2^2(e_{12})$. If we do so, we are then faced with the

dilemma as to what to do if we get an estimate of the precision of a system by the earlier methods and by the method of this section and the estimates do not agree. Recall now from (3.6) and (6.5) that the variance of the error in question is bounded above by the minimum of the set of variances of differences of the random variables in question. That is:

$$\sigma^2(e_a) \leq \min\{\sigma^2(X_a \pm X_{a'}), a \neq a' = 1, \ldots, k\}, a = 1, \ldots, k \quad (6.7)$$

and after adding

$$\sigma^2(e_{ij}) = \sigma^2(\xi_i) + \sigma^2(\xi_j) \leq$$

$$\min\{\sigma^2(X_{ik} \pm X_{i'k'}) + \sigma^2(X_{jl} \pm X_{j'l'}), \ i' \neq i; \ j' \neq j; \ i \neq k/i'; \quad (6.8)$$

$i \neq k' \neq i'; \ j \neq l \neq j'; \ j \neq l' \neq j'; \ i, i', j, j', k, k', l, l' = 1, \ldots, m\}$.

Quite obviously, if system a consists of sensors i and j, then:

$$\sigma^2(e_a) = \sigma^2(e_{ij}) = \min\{[\text{RHS of } (6.7)], [\text{RHS of } (6.8)]\}. \quad (6.9)$$

Hence, we take

$$\hat{\sigma}^2(e_a) = \hat{\sigma}^2(e_{ij}) = \min\{\hat{\sigma}_1^2(e_a), \hat{\sigma}_2^2(e_{ij})\} \text{ where sensors i and j}$$

comprise system a. $\quad (6.10)$

7. <u>Comprehensive Example</u>: In the following example, there are four sensors, labeled 1 through 4 used in such a way that each of the

$\binom{4}{2} = 6$ pairs labeled A through F, yields a velocity. In addition

there are two other systems, labeled G and H that are independent of each other and systems A through F that each yield a velocity. The systems A through F are obviously not independent, however, certain pairs are. The labels in the table to follow show that system A is independent of system B, G and H, but dependent on C, D, E, and F. Similar statements apply to systems B through F.

System	A	B	C	D	E	F	G	H
Sensor	1,2	3,4	1,4	2,3	1,3	2,4	5,6	7,8
	5.108	4.364	4.408	3.877	4.975	4.607	4.682	4.348
	4.649	2.784	3.422	2.967	4.289	3.360	5.957	5.727
	0.751	0.078	0.238	0.011	0.865	0.529	0.170	0.080

The following results are obtained.

$$s^2(x_A - x_B) = 0.447 \qquad s^2(x_D - x_G) = 2.201$$

$$s^2(x_A - x_G) = 1.100 \qquad s^2(x_D - x_H) = 1.538$$

$$s^2(x_A - x_H) = 1.074 \qquad s^2(x_E - x_F) = 0.111$$

$$s^2(x_B - x_G) = 2.949 \qquad s^2(x_E - x_G) = 1.598$$

$$s^2(x_B - x_H) = 2.901 \qquad s^2(x_E - x_H) = 1.538$$

$$s^2(x_C - x_D) = 0.025 \qquad s^2(x_F - x_G) = 2.548$$

$$s^2(x_C - x_G) = 2.001 \qquad s^2(x_F - x_H) = 2.477$$

$$s^2(x_C - x_H) = 1.945 \qquad s^2(x_G - x_H) = 0.015$$

These value and (3.7) yield the following:

$$\hat{\sigma}_1^2(e_A) = 0.447 \qquad \hat{\sigma}_1^2(e_E) = 0.111$$

$$\hat{\sigma}_1^2(e_B) = 0.447 \qquad \hat{\sigma}_1^2(e_F) = 0.111$$

$$\hat{\sigma}_1^2(e_C) = 0.025 \qquad \hat{\sigma}_1^2(e_G) = 0.015$$

$$\hat{\sigma}_1^2(e_D) = 0.025 \qquad \hat{\sigma}_1^2(e_H) = 0.015$$

Values for the dependent systems are obtained as follows.

$$s^2(x_A - x_C) = 0.137 \qquad s^2(x_B - x_E) = 0.224$$

$$s^2(x_A - x_D) = 0.222 \qquad s^2(x_B - x_F) = 0.028$$

$$s^2(x_A - x_E) = 0.056 \qquad s^2(x_C - x_E) = 0.025$$

$$s^2(x_A - x_F) = 0.306 \qquad s^2(x_C - x_F) = 0.034$$

$$s^2(x_B - x_C) = 0.099 \qquad s^2(x_D - x_E) = 0.055$$

$$s^2(x_B - x_D) = 0.115 \qquad s^2(x_D - x_F) = 0.029$$

From (5.1) we obtain:

$$\hat{\sigma}_1^2(e_A) = 0.279 \qquad \hat{\sigma}_1^2(e_D) = 0.253$$

$$\hat{\sigma}_1^2(e_B) = 0.223 \qquad \hat{\sigma}_1^2(e_E) = 0.100$$

$$\hat{\sigma}_1^2(e_C) = 0.241 \qquad \hat{\sigma}_1^2(e_F) = 0.108$$

Note that these values for $\hat{\sigma}_1^2(e_A)$, $\hat{\sigma}_1^2(e_B)$, $\hat{\sigma}_1^2(e_E)$ and $\hat{\sigma}_1^2(e_F)$ are

lower than those obtained previously.

By way of comparison, the Grubbs estimates on this data are:

$$\hat{\sigma}_C^2(e_A) = -0.104 \qquad \hat{\sigma}_C^2(e_E) = 0.220$$

$$\hat{\sigma}_C^2(e_B) = 1.734 \qquad \hat{\sigma}_C^2(e_F) = 1.187$$

$$\hat{\sigma}_C^2(e_C) = 0.698 \qquad \hat{\sigma}_C^2(e_G) = 1.486$$

$$\hat{\sigma}_C^2(e_D) = 0.595 \qquad \hat{\sigma}_C^2(e_H) = 1.332$$

Further, from (3.7) and (6.2) we obtain the following estimates for the precision of each of the sensors.

$$\hat{\sigma}_2^2(\xi_1) = 0.025 \qquad \hat{\sigma}_2^2(\xi_5) = 0.015$$

$$\hat{\sigma}_2^2(\xi_2) = 0.025 \qquad \hat{\sigma}_2^2(\xi_6) = 0.015$$

$$\hat{\sigma}_2^2(\xi_3) = 0.025 \qquad \hat{\sigma}_2^2(\xi_7) = 0.015$$

$$\hat{\sigma}_2^2(\sigma_4) = 0.025 \qquad \hat{\sigma}_2^2(\xi_8) = 0.015$$

Finally, from (6.10)

$$\hat{\sigma}^2(e_A) = 0.050 \qquad \hat{\sigma}^2(e_E) = 0.050$$

$$\hat{\sigma}^2(e_B) = 0.050 \qquad \hat{\sigma}^2(e_F) = 0.050$$

$$\hat{\sigma}^2(e_C) = 0.025 \qquad \hat{\sigma}^2(e_G) = 0.015$$

$$\hat{\sigma}^2(e_D) = 0.025 \qquad \hat{\sigma}^2(e_H) = 0.015$$

8. **Bias:** Although there is the possibility of creating a furor over semantics, it appears to this author that one can only estimate the relative bias between instruments. This is to say that in reality there is only one model where bias is estimable. Suppose instrument 1 is chosen as the "standard", then modify (2.1) as

and
$$X_{1j} = T_{1j} + e_{1j}$$

Now
$$X_{ij} = T_{ij} + E_{ij} + e_{ij} \quad , \quad i = 2, \ldots, k; \quad j = 1, \ldots, n.$$

and
$$\mu(X_1) = \mu(T_1)$$

$$\mu(X_i) = \mu(T_i) + \mu(E_i), \quad i = 2, \ldots, k.$$

If in addition, one assumes that $T_{1j} = T_{ij}$, $i = 2, \ldots, k$, i.e. each

instrument is measuring the same phenomenon on each of the n events.
Then
$$X_{1j} - X_{ij} = -E_{ij} + e_{1j} - e_{ij} \quad , \quad i \neq 1, \quad j = 1, \ldots, n$$
and

$$\mu(X_1 - X_i) = -\mu(E_i) \quad , \quad i \neq 1,$$
or

$$\mu(E_i) = \mu(X_i - X_1), \quad i \neq 1.$$

So the bias of instrument i relative to instrument 1 is estimable. Note that this model does not require equality of the n values T_{11}, $T_{12}, \ldots, T_{1n}$ (or $T_{i1}, T_{i2}, \ldots, T_{in}$). Therefore, this is an acceptable model in the ballistic situation where the velocity differs from round to round. However, it does require the various systems to sense the velocities over the same interval to assure $T_{1j} = T_{2j} = \ldots = T_{kj}$.

The instance of measuring a gage block is a special case of the above where the standard instrument is used to determine the length, mass, etc. of the block, then the relative difference of measurement by each instrument is observed. This difference is the bias of the test instrument relative to the standard. If one considers the standard to be absolute, then one would call this the absolute bias of the instrument.

9. <u>Conclusions</u>: Estimates of measurement precision for systems or individual sensors can be derived from sample data in the most general context. No assumptions are required on the relationships between the true event values, nor on the nonrandom errors, nor on whether the nonrandom errors are dependent on the true event values. Further, no distributional assumptions other than symmetry about zero must be made regarding the random errors. Finally, the only assumptions that are necessary are that the random errors are independent of the true event values and of the nonrandom errors and that at least one system is independent of the others.

In every example analyzed by the author, the estimates of precision calculated by this method were lower than the absolute value of the point estimate obtained from Grubbs' method, even when the assumptions of Grubbs' method were reasonable. Further, one is never faced with the dilemma of how to handle a negative estimate for a variance.

In those cases where an absolute standard does not exist, but the k instruments are measuring the same event, the relative bias of k-1 of the instruments can be estimated relative to the instrument chosen as the standard. This method does not require that the n true values read with each instrument be the same, only that the k instruments each measure the same n events.

References

Churchill Eisenhart, "Realistic Evaluation of The Precision and Accuracy of Instrument Calibration Systems", Journal of Research of the National Bureau of Standards, Vol 67C, No. 2, April–June 1963.

Frank E. Grubbs, "On Estimating Precision of Measuring Instruments and Product Variability", Journal of the American Statistical Association, Vol. 43, June 1948, 243–64.

Frank E. Grubbs, "Errors of Measurement, Precision, Accuracy and the Statistical Comparison of Measuring Instruments", Technometrics, Vol. 15, No. 1, Feb 1973, pp 53–66.

G. J. Hahn and W. Nelson, "A Problem in the Statistical Comparison of Measuring Devices", Technometrics, Vol. 12, No. 1, 1970, pp 95–102.

R. Choudary Hanumara, "Estimating Imprecisions of Measuring Instruments", Technometrics, Vol. 17, No. 3, August 1975, pp 299–302.

R. Choudary Hanumara, "Estimating Variances in Simultaneous Measurement Procedures", Letter to the Editor of the American Statistician, Vol. 29, No. 2, May 1975, pp 108–9.

John L. Jaech, "Large Sample Tests of Grubbs' Estimators of Instrument Precision with More Than Two Instruments", Technometrics, Vol. 18, No. 2, May 1976, pp 127–33.

C. J. Maloney and S. C. Rastogi, "Significance tests for Grubbs' Estimators", Biometrics, Vol. 26, 1970, pp. 671–6.

G.K. Shukla, "Some Exact Tests of Hypothesis About Grubbs' Estimators", Biometrics, Vol. 29, 1973, pp 373–77.

W. A. Thompson, Jr., "The Problem of Negative Estimates of Variance Components", The Annals of Mathematical Statistics, Vol. 33, No. 1, March 1962, pp 273–89.

W. A. Thompson, Jr., "Precision of Simultaneous Measurement Procedures", Journal of the American Statistical Association, Vol, 58, No. 302, June 1963, pp 474–9.

US Department of Commerce, National Bureau of Standards, "Precision Measurement and Calibration (Statistical Concepts and Procedures)", Special Publication 300, Vol. 1, February 1969.

LIQUID PROPELLANT DISTRIBUTION AND COMBUSTION

*RAO YALAMANCHILI, DR.
TECHNOLOGY BRANCH, ARMAMENT DIVISION
FIRE CONTROL & SMALL CALIBER WEAPON SYSTEMS LABORATORY
U.S. ARMY ARMAMENT RESEARCH AND DEVELOPMENT CENTER, DOVER, NJ   07801

I.   INTRODUCTION

Even though the United States was the first to introduce the concept of a liquid propellant gun more than three decades ago, the United States may not be the first to introduce a fielded weapon unless otherwise drastic measures are undertaken to reverse the current situation by providing more funds and manpower immediately. The extent of Soviet involvement in liquid propellant guns is unknown. We should not wait for another surprise just like the one of launching Sputnik, the artificial satellite, in 1957. If that does happen, we may wind up with another crash program of uneconomical layout to regain the technology lead.

Liquid propellants have many potential advantages over solid propellants for use in guns, including elimination of cartridge cases, reduced gun tube erosion, and higher muzzle velocities. One can also view the advantages as elimination of problems associated with solid propellants. For example, the products of combustion of solid propellants may contain about 20% (much higher than its liquid counterpart) of fuel rich carbon compounds and thus there is less muzzle flash in the case of liquid propellants. The liquid propellants are not high explosives when compared to nitrocellulose and nitroglycerine and thus are less vulnerable to enemy fire or accidents. The impetus, or propellant force per unit mass, is high for liquid propellants because of high heat of explosion and lower molecular weight of products of combustion. The biggest contributor for erosion is adiabatic flame temperature which is less for liquid propellants than solid propellants for a given specific force. The liquid propellants can be stored outside of the self-propelled howitzers and tanks to augment the limited space inside. In addition, liquid propellants possess higher loading densities than solid propellants and thus require less space and contribute to improvement in logistics. Critics say there exists possibility of a variable burning surface area for liquid propel-

lants, especially in bulk-loaded guns. What about possibility of fractured solid propellant grains due to excessive stresses that causes a catastrophic increase in burning surface area?

Basically, there are three types of liquid propellant gun systems. There is a bipropellant system, which is more complex than monopropellant systems. Monopropellant systems studied have been either of the bulk loaded or regenerative type. Most of the work until now has been concentrated on the bulk loaded liquid propellant system. Here, the complete charge is introduced at one time in the combustion chamber and allowed to remain in contact with the combustion zone. The state-of-the-art is very limited in other regenerative liquid propellant systems. First the propellant is loaded into a reservoir and later it is injected into the combustion chamber as needed. At present, the research is taking place on both systems. However, this paper is limited to only the regenerative liquid propellant system.

II. PHYSICAL INSIGHT

The schematic of a regenerative liquid propellant system is shown in Figure 1. A monopropellant which contains both fuel and oxidant is considered because of simplicity. A differential piston is used, i.e., the area of a piston facing the projectile is much larger than the area of a piston facing the propellant. Therefore, there is a net force acting in a rearward direction even if the pressure is the same on both sides. A small amount of monopropellant is injected into the region between the bolt and the piston head and the inlet and orifices which are located in the piston head are sealed off. An electric primer is introduced into the combustion chamber between the piston head and the projectile. The primer gases push the piston rearward. The liquid propellant is compressed and higher liquid pressures are generated. This high pressure liquid propellant is injected into the combustion chamber through the injectors. Some of the injected propellant, which is in the form of a spray, is burned and higher pressures are generated. The process continues until the propellant is exhausted. The objective of this paper is to develop mathematical models and solution techniques in order to obtain detailed interior ballistics. These models are based on firings of the General Electric regenerative test fixture and neutron radiographic data.

It is essential to validate the mathematical model in order to obtain a workable and optimum regenerative liquid propellant gun system model. In preparation for constructing such a model a unique experimental technique, high speed neutron radiography, has been utilized to observe the ignition and combustion of liquid propellants in the regenerative piston configuration. Neutrons generated in a TRIGA reactor were utilized to observe the liquid propellant during the firing cycle. Neutrons have the

advantage of being attenuated by hydrogen and hydrogen-containing materials, such as liquid propellants. They pass through metals relatively undisturbed. Utilization of this technique allowed for obtaining high speed motion pictures of the entire regenerative liquid propellant ballistic process.

The regenerative piston films have undergone image enhancement at Los Alamos National Laboratory and interior ballistic data such as piston location, jet width, surface area of jet, volume of jet, and also proportional distribution of liquid propellant mass was obtained. A .35 caliber test fixture was utilized. Separate tests provided liquid propellant pressure inside the resorvoir, piston position, and also chamber pressure. However, caution is to be exercised in relating the two different sets of data because these are obtained independently. Some of the data, as it is, did not appear reproducible. However, rearranging the data with a common base yielded reproducible results within 10% variation.

Initially, the liquid pressure is lower than the chamber pressure, even though the piston is moving rearward. It takes quite a while for the liquid pressure to exceed the chamber pressure. If a compressible model is utilized, it may be implied that there is not injection of propellant until that time because of Bernoulli's equation. However, the neutron radiographic pictures indicate the existence of a jet earlier than this time:

TABLE 1. EXISTANCE OF JET/SPRAY

STATUS	DISPLACEMENT OF PISTON, INCHES	
	SHOT #19	SHOT #11
No jet/spray	0.037	0.025
Jet/spray exists	0.067	0.054

No classical theories are reasonable in this situation because both the compression wave and the pressure-relief orifice are present at the same location. Thus, the net compression of the liquid is small and there may not be much significant change in the space-mean liquid pressure, especially during the initial portion of the ballistic cycle. According to theory, the jet/spray is not possible without appreciable increase in pressure. However, the neutron radiographic data indicate the existance of the orifice at the same location. Until a satisfactory theory is developed for such a configuration, an incompressible assumption has been used in the initial period of the ballistic cycle and later compressibility is considered as the pressure increases.

The physical phenomena in a regenerative liquid propellant system is quite complex. The problem is an unsteady, two-phase and three-dimensional flow in a highly reactive environment. This is further complicated by the unknown moving boundary conditions. The kinetics of the propellant decomposition is currently unknown. The regenerative liquid propellant problem is complex due to an unsteady environment and interaction between numerous complex droplets of varying sizes and shapes. Therefore, a practical minded approach is considered as a first cut. There is plenty of scope for future improvement in the modeling process. A transient one-dimensional approach fits this category.

III. PROPELLANT DISTRIBUTION

The OTTO Fuel II is considered. High speed neutron radiographs were taken at Oregon State University from the firings of a .35 caliber in-line regenerative piston fixture designed by the General Electric Company. The films were subjected to image enhancement and data reduction techniques at Los Alamos National Laboratory. The typical fuel properties are as follows: density of propellant (ρ_{ℓ}) = 0.0567 lb/in^3, surface tension (σ) = 0.002361 lbf/ft, kinematic viscosity $(\nu$ = 0.0000384 ft^2/sec, impetus (F) = 270,000 ft lbf/lbm, bulk modulus (β) = 300,000 psi, adiabatic flame temperature (T_0) = 3600 °R, thermal conductivity = 0.202 BTU/hr-ft-°R, specific heat (C_p) = 0.445 BTU/lbm - °F.

The jet and spray characteristics will depend upon the orifice design, Weber number, Reynolds number, and the density ratio between the hot gases and liquid propellant. These are computed based upon typical operating conditions: Length to diameter ratio = 4, maximum gas density (ρ), = 0.0103 lb/in^3 (computed at maximum pressure and available volume between the piston and the projectile including corrections due to co-volume (η), and unburnt propellant volume), maximum liquid propellant injection velocity = 850 ft/sec (for a pressure difference of 6000 psi). The Reynolds number varies between 0 and 148,000, whereas the Weber number varies between 0 and 3,000,000. The maximum gas to liquid density ratio becomes 0.23.

The flow is essentially inviscid close to the orifice exit plane; however, a shear layer starts developing with an essentially zero thickness as the flow proceeds downstream. The inner boundary of this region cuts the axis of the jet at a certain distance from the nozzle exit plane. From here onward the flow will be fully viscous and turbulent in nature. The viscosity of the fluid plays the major role in obtaining the characteristics of the flow field.

Even if the surrounding gas is calm especially near the orifice exit plane because of the zero velocity requirement at the breech end, the

ejecting action of the jet and the turbulence of the jet flow field causes surrounding gas to mix with the jet flow field. The mass entrained through the shear layer between any two cross sections of the flow field can be obtained by calculating the total mass at the two cross sections. The density distribution and the velocity profile at any cross section should be known in order to obtain the mass entrainment.

Limited data reduction of density distribution inside the neutron radiographs has been performed based on data provided by Los Alamos national Laboratory. The results indicate an exponential pattern inside the jet/spray implying the possibility of diffusion of a jet into a form of spray. There is no data regarding the velocity distribution across the jet. An existing literature revealed a power-law profile in similar situations. If such a power law is invoked for velocity distribution, the mass entrainment can be approximated.

An examination of neturon radiographic pictures indicate that the jet is of straight and cylindrical form and slightly larger than the diameter of the orifice. Professor Bracco of Princeton University did dextensive research for a decade in gasoline and diesel injection. Extension of those results to the regenerative liquid propellant system should show a divergent spray with an angle of 15° initially when the density of gas is low and about 18° whenever the gas density reaches its peak. The reason why such a behavior is not seen in neutron radiographs is due to the nature of the experimenal technique, i.e., severe hydrogen concentration in the center of the spray and mild in the outer region due to entrainment phenomena, and also vaporization and combustion of droplets near the surrounding hot propellant gases. Therefore, neutrons attenuated by hydrogen and hydrogen-containing materials produce such an illusion on neutron radiographs.

However, there is something else of concern here. The jet is straight and appears too long. This is not breaking up like gasoline and diesel injection sprays in internal combustion engines. The design of regenerative pistons and further testing is imperative if an optimum liquid propellant gun system is desired.

In the model it is assumed that a primer of known charge (C_i), force (F_i), and adiabatic flame temperature (T_i) is present in the combustion chamber, between the piston head and the projectile base. On electric ignition, the primer produces pressure which moves the piston into the propellant reservoir. The liquid monopropellant inside the reservoir (between the bolt and the piston head) is compressed which, in turn, generates higher pressures. The pressure difference across the piston head causes an injection of liquid propellant into the combustion chamber (hot gas combustion zone) through orifices. If an incompressible assump-

tion for liquid monopropellant is invoked, the rate of propellant discharge (M) can be represented by:

$$\frac{dM}{dt} = \rho_\ell (A_z - A_R) \frac{dz}{dt}$$ (1)

Where A_z = cross sectional area of piston head, A_R = cross sectional area of piston rod, z = position of piston, and t = time.

This equation is based on the condition that the volume displaced by the piston must be equal to the volume of propellant discharged into the combustion chamber.

It is assumed that the pressures inside the liquid propellant gun are similar to those observed in a conventional gun. An incompressible liquid assumption may not be a good idea at higher pressures. The following equations may be derived for a compressible fluid:

$$\frac{dM}{dt} = C_d \rho_\ell A_h V_\ell$$ (2)

$$V_\ell = \sqrt{\frac{2g|P_\ell - P_z|}{\rho_\ell}}$$ (3)

Where C_d = discharge coefficient, A_h = area of the orifice, V_ℓ = velocity of the orifice, g = gravitational constant, P_ℓ = pressure in the reservoir, and P_z = pressure in combustion chamber near the piston. The discharge coefficient can be a function of Reynold's number.

The piston motion may be described by Newton's second law of motion. The forces acting on the piston in the axial direction are pressure forces and the thrust produced by the action of a jet/spray. The resulting equation for piston motion becomes:

$$\frac{W_p}{g} \frac{dV_z}{dt} = P_z A_z - P_\ell (A_z - A_R) + \rho_\ell A_h V_\ell^2 / g$$ (4)

$$\frac{dz}{dt} = V_z$$ (5)

Where W_p = weight of piston and V_z = velocity of piston.

The liquid propellant pressure variation with respect to chamber pressure may be expressed mathematically by the use of the definition of bulk modulus (β) and control volume concepts:

$$\frac{dP_\ell}{dt} = \frac{\beta}{U_\ell} [(A_z - A_R)V_z - C_d A_h V_\ell] \tag{6}$$

The ratio of change of volume of liquid propellant (U_ℓ) inside the reservoir can be expressed as

$$\frac{dU_\ell}{dt} = (A_z - A_R)V_z \tag{7}$$

The mathematical definition of bulk modulus of a liquid may be stated as

$$\beta = -U_\ell \frac{\partial P_\ell}{\partial U\ell} \tag{8}$$

The bulk modulus is approximated as

$$\beta = a + bP\ell \tag{9}$$

where a and b are constants determined experimentally.

The rate of change of liquid propellant density, based on the definition of density, is formulated as

$$\frac{d\rho_\ell}{dt} = \frac{(M_o - M)(A_z - A_R)}{[U_o - (A_z - A_R)V_z]^2} V_z - \frac{\rho_\ell C_d A_h V_\ell}{U_o - (A_z - A_R)z} \tag{10}$$

Where M_o = initial amount of liquid propellant in the reservoir, and U_o = initial volume of liquid propellant reservoir.

There is no satisfactory theory for droplet size prediction and it is also impossible to obtain uniform droplets even in any chosen time incre-

ment. Therefore, an empirical derivation, based on Adelberg, is utilized in the initial portion of the firing cycle.

$$D_\ell = CD_h \tag{11}$$

where D_ℓ = mean diameter of liquid propellant droplet, C = empircal constant (function of orifice and propellant properties), and D_h = diameter of orifice. It is not uncommon to utilize the Weber number criteria for droplet size prediction. The Weber number in terms of relative velocity (ΔV) between gas and liquid is defined as

$$We = \frac{\rho(\Delta V)^2 D_\ell}{2\sigma} \tag{12}$$

The velocity of droplets at the orifice exit is given by equation 3. The vaporization of droplets in the spray is dependent on the ballistics and trajectories of individual drops, or groups of droplets, which may be determined from the empirically derived drag coefficients, C_D, using an expression of the form

$$\frac{dV_\ell}{dt} = 0.75 \frac{C_D \rho \Delta V |\Delta V|}{\rho_\ell D_\ell} \tag{13}$$

The correlations which allow both for liquid evaporation and for droplet distortion at high relative velocities for burning and nonburning fuel droplets are as follows:

$$\begin{aligned} C_D &= \ 27\ Re^{-0.84}, & 0 < Re < 80 \\ &= 0.271\ Re^{0.217}, & 80 < Re < 10^4 \\ &= 2, & Re > 10^4 \end{aligned} \tag{14}$$

The integration of equation 7 yields the velocity of droplets as a function of the time based on aerodynamic behavior in a reactive environment. Then, the definition of velocity ($V_{\ell} = dX_D/dt$) and its integration can provide the information regarding the location of droplets (X_D) as a function of time. The calculations indicate that the propellant droplets can reach the base of the projectile in the initial portion of the ballistic cycle.

The number of droplets is determined from the mass of liquid propellant injected and the mass of a single droplet based on droplet size predictions. A liquid drop removed from a stream or jet may be exposed to the aerodynamic pressure effect of a high relative velocity. If the pressure is large enough to overcome the restoring force of the drop surface tension, the drop will disintegrate into smaller droplets. The Weber number is considered as an effective indicator of the necessary gas dynamic environment for secondary break-up. The occurrence of such an event is considered likely if the Weber number is greater than 6 and 10 for low and high viscosity fluids, respectively. The droplet break-up time is calculated by use of an existing literature and found to be negligible for typical gun conditions.

IV. VAPORIZATION AND COMBUSTION

It is desirable to know the required droplet sizes for rapid combustion and also for a meaningful match between analytical interior ballistic results and experimental results. Consider the vaporization process of an individual droplet. This is shown in Figure 2. The time to burn out, t_b, is of the order of 1 millisecond for some droplets. Since this is of the same order of magnitude as an interior ballistic cycle, it is better to study this combustion in detail. The sequence of combustion of a monopropellant droplet may be divided into two phases: (1) heat up phase, and (2) combustion phase. The droplet is heated up by forced convection and radiation. The heat-up phase ends as the drop nears its boiling temperature. During combustion, the drop remains at a wet bulb temperature (T_{wb}) which is slightly below its boiling point. The droplet size remains almost the same until the end of the heat-up phase and drops off rapidly during the vaporization and combustion phase.

When the system pressure increases, the boiling and wet bulb temperatures increases, which tends to increase the heat-up time. However, the burning rate increases as well, so that the total time for combustion decreases for monopropellants as the pressure is increased. With increasing pressure, the drop eventually reaches its thermodynamic critical point and gasifies during heat-up. In this case, the combustion reaction takes place completely in the gas phase, in a process similar to that occurring in a premixed turbulent gas flame. This is shown in Figure 3.

Faeth estimated heat-up times by assuming a constant convective heat transfer coefficient (h) between the spherical droplet and products of combustion:

$$\frac{hD_\ell}{K_g} = 2R \tag{15}$$

Where K_g is the thermal conductivity of the gases and R is a factor of order of magnitude unity which also depends upon the Reynolds number of the droplet. A simple heat conduction model with convection yields

$$t_{hu} = - \frac{\rho_\ell C_p D_\ell^2}{12 K_g R} \ell n \left(\frac{T_{wb} - T_\infty}{T_o - T_\infty} \right) \tag{16}$$

T_o is the initial temperature of liquid and T_∞ is the gas temperature. The following table is obtained by the use of typical OTTO Fuel prperties: $R = 1$, $T_{wb} = 220°C$, $T_o = 20°C$, $T_\infty = 1730°C$.

TABLE 2. HEAT-UP TIMES OF DROPS

D_ℓ (microns)	t_{hu} (millisec)
10	0.027
20	0.109
40	0.436
80	1.744

The droplet size may be on the order of 20 to 40 microns to yield heat-up times in the range 0.1 - 0.4 msec in comparison to an interior ballistic cycle in the range of 3-9 msec.

The second phase involves vaporization and chemical reactions (combustion). The burning rate equation similar to solid propellants is assumed, but with different coefficients and exponents for Weber number less than the critical Weber number:

$$\frac{dm}{dt} = \rho_\ell S \frac{d\delta}{dt} \quad ; \quad \frac{d\delta}{dt} = BP^n \tag{17}$$

Where m = mass of propellant converted to gas, S = total surface area of all droplets, δ = distance burned normal to the surface, P = mean chamber pressure, B = 0.04, n = 0.75 for 3.5 MP < P < 17.2 MP, B = 0.011, n = 1.27 for 52 MP < P < 207 MP.

An estimate of time of combustion $(t_c = t_b - t_{hu})$ may be obtained by the above burning rates and realization that 90% of the mass of a droplet is burned by the time the radius is changed to 50%. Anyhow, the linear burning laws don't hold good beyond this limit. Thus, the combustion time is approximated as

$$t_c = (D_\ell/4)/(d\delta/dt) \tag{18}$$

Table 3 provides combustion times for typical gun barrel conditions.

TABLE 3. COMBUSTION TIMES OF DROPLETS (MSECS)

D_ℓ (microns)	PRESSURE MPa				
	5	10	50	100	200
10	1.86	0.83	0.015	0.006	0.003
20	3.72	1.66	0.030	0.013	0.005
40	7.44	3.32	0.060	0.025	0.011
80	14.88	6.64	0.120	0.050	0.022

The droplet size effect on burning time is linear. At low pressures, the large droplets have delay times of the same order of magnitude as interior ballistic cycle time. At high pressures, the combustion time is short and t_{hu} will dominate the delay between injection and reaction. By combining the above results, one can formulate an equation of the following form for ignition delay time (t_d):

$$t_d = a + bP^{-c} \tag{19}$$

where a and b are functions of injector characteristics since these depend on droplet size. Such an equation is useful for matching analytical results against experimental data. If the Weber number is greater than the critical Weber number, then the number and size of existing droplets

are modified and the newly injected droplets are small and burned immediately.

The free volume available in the chamber for gas molecules can be calculated based on control volume concepts. The free volume (U) is defined by equation 20.

$$U = U_0 + A_z z + Ay \frac{M - m}{\rho_\ell} - \eta m \qquad (20)$$

$$\frac{dU}{dt} = A_z \frac{dz}{dt} + A\frac{dy}{dt} + \frac{M - m}{\rho_\ell^2}\frac{d\rho_\ell}{dt} - \frac{1}{\rho_\ell}(\frac{dM}{dt} - \frac{dm}{dt}) - \eta\frac{dm}{dt} \qquad (21)$$

Where A = bore cross sectional area and y = location of projectile.

The relation for propellant gas density (ρ) can be developed based on the definition of density. That is

$$\frac{d\rho}{dt} = \frac{1}{U}\frac{dm}{dt} - \frac{m}{U^2}\frac{dU}{dt} \qquad (22)$$

The equation of state can be used to determine the mean pressure (P) in the combustion chamber

$$P = \frac{mF}{U}\frac{T}{T_0} + \frac{C_i F_i}{U}\frac{T}{T_i} \quad \text{or} \qquad (23)$$

$$\frac{dP}{dt} = \frac{F}{T_0 U^2}[UT\frac{dm}{dt} + Um\frac{dT}{dt}\ mT\frac{dU}{dt}] + \frac{F_i C_i}{T_i}[\frac{1}{U}\frac{dT}{dt} - \frac{T}{U^2}\frac{dU}{dt}] \qquad (24)$$

The propellant gas mean temperature (T) can be determined from the conservation of energy equation which can be stated as the energy released by the burning fuel equals the sum of the following quantities: internal energy of propellant gases, kinetic energy of propellant gases, kinetic energy of projectile, work done on piston, heat lost to gun tube and energy loss due to projectile friction.

$$\frac{dT}{dt} = \{\frac{\varepsilon - 1}{\gamma - 1}(mF + C_i F_I) + \int F_z dz + KE_{gas} + (1 + \theta)\frac{W_B}{2g}V_B^2\}.$$

(25)

$$\cdot \frac{F dm/dt}{(\gamma - 1)T_0[\frac{mF}{(\gamma - 1)T_0} + \frac{C_i F_i}{(\gamma - 1)T_i}]^2} + \{\frac{1 - \varepsilon}{\gamma - 1}F\frac{dm}{dt} - \frac{d}{dt} \int F_z dz$$

$$- \frac{d}{dt}(KE_{gas}) - (1 + \theta)\frac{W_B}{g} V_B\frac{dV_B}{dt}\}/[\frac{mF}{(\gamma - 1)T_0} + \frac{C_i F_i}{(\gamma - 1)T_i}]$$

Where $F_z = P_z A_z + \rho_g A_h V_g^2/g$, ε = ratio of heat loss to the tube and energy released by burned propellant, θ = ratio of energy lost due to projectile friction and kinetic energy of projectile, KE_{gas} = kinetic energy of propellant gases, W_B = weight of the projectile, and γ = ratio of specific heats.

Finally, the projectile motion can be represented based on Newton's second law of motion:

$$\frac{dV_B}{dt} = \frac{Ag}{W_B} (P_s - P_r) \text{ and } \frac{dy}{dt} = V_B$$

(26)

Where V_B = velocity of projectile, P_s = gas pressure at the base of the projectile, and P_r = resistive pressure due to friction between projectile and tube and also due to air resistance ahead of the projectile.

V. VALIDATION

A mathematical model is constructed to represent various physical processes inside a regenerative liquid propellant gun. The model is based on sound engineering judgments from existing literature, state-of-the-art neutron radiographs, and firing data. To complete the model, it is assumed that there is a linear gas velocity gradient between piston and projectile. This is much better than the uniform density approximation used by many others. The gas density is allowed to vary linearly between piston and projectile. The transient one-dimensional (partial differential) equation of motion is integrated with appropriate boundary condi-

tions to yield the pressures at the base of the projectile and also at the piston head. In all, there are 15 coupled ordinary differential equations which are solved by numerical integration techniques due to lack of possibility of closed-form solutions. The unknowns are M, m, ρ_ℓ, U_ℓ, V_z, z, P, T, ρ, V_B, y, P_z, P_s, X_D, δ, and U.

Consider the motion of the piston. The fluid pressures and the areas subjected to those pressures are different. The piston is in dynamic equilibrium during a portion of the ballistic cycle. One can set up a force balance and obtain a relationship between fluid pressures and ratio of areas. The ratio of liquid propellant pressure to chamber pressure attained a value of 25/21. This is the same as the geometrical (area) ratio whenever the piston is in dynamic equilibrium.

The trial runs indicate that there is a difference in pressure between analytical and experimental results; however, the trends are similar. Therefore, the differences can be eliminated by altering some features of the model including ignition delay or incubation period and other constants. The typical sample problem is shown: A_h = 0.0314, C_i = 0.000243, F_i = 3,298,969, T_i = 3486°K, t_d = 0.002, M_o = 0.0136, W_B = 0.022, T_o = 2000°K, A_B = 0.0483, A = 0.0962, A_z = 0.305, C_d = 0.9, η = 30, $P_s(0)$ = 6000, ρ = 0.0445, θ = 0.05, W_p = 0.1108, P_r = 3000, γ = 1.26, V_o = 0.839, F = 3,377,616, and z_{max} = 1.2.

The liquid pressure and gas pressure are shown in Figure 4. This may not be the typical case because the amount of liquid propellant available (5CC) is consumed at the time slightly past the peak pressures. Therefore, the flat section (a typical characteristic of liquid propellants) of the pressure-time curve cannot be seen. Anyone might notice this desirable property if more propellant and longer piston stroke are made available. There is a significant ignition delay. The gas pressure builds up and compresses the liquid. Therefore, the gas pressure is higher than the liquid pressure initially. There is about a 0.6 millisecond time elapse for the liquid pressure to exceed the gas pressure. The agreement between theory and experiment is not that good because this is the first time a comparison has been made and no attempts are made to fine tune the model. There are quite a number of constants, coefficients, and exponents that take flexible values and it doesn't not make sense to adjust them until more test cases are generated. Reasonable agreement is obtained between theory and experiment for piston motion. The experimental muzzle velocity is 2623 ft/sec, whereas the first cut theoretical attempt yields 2670 ft/sec.

VI. PARAMETRIC ANALYSIS

A weapon system or concept may not achieve optimum operation with cut and try techniques. However, a validated mathematical model can yield an optimum design if a parametric analysis is performed. A number of parameters are incorporated into the model. Some of these parameters are varied to show their influence on overall operation of the system. The trends will remain valid. The actual values may vary after validation with more experimental data.

The coefficient of discharge is an important parameter. The state-of-the-art is limited. No one can state exactly the value for any orifice and transient flow conditions. Further basic research is needed to relate to orifice design and flow conditions (Reynolds number). Table 4 shows some of the results (maximum values) for typical values of coefficient of discharge.

TABLE 4. EFFECT OF COEFFICIENT OF DISCHARGE

C_d	P_ℓ (KPSI)	P (KPSI)	V_B (FPS)	V_z (FPS)
0.6	18.6	14.7	1930	30.7
0.7	21.4	16.8	2060	39.8
0.8	24.3	19.0	2160	49.9
0.9	26.1	20.4	2190	59.6
1.0	29.0	22.6	2210	71.2

As coefficient of discharge increases, the peak values of liquid pressure, gas pressure, velocity of projectile, and velocity of piston will increase. However, the effect at the high end is not as dramatic as at the low end. All values of C_d, listed here, are probable, depending upon the design and operating conditions. There are some investigators, in the past, who used a C_d as high as 1.4. Such a value is physically absurd from its definition. They apparently did not care as long as they got some agreement between model and experiment. The peak piston velocity is significant. Some mechanism may be required to reduce the speed to zero.

The typical initial chamber volume is varied and the peak values of the same four quantities are tabulated in Table 5. As the initial chamber volume is increased, the maximum values are reduced, thus affecting the performance of the weapon system. The variations are similar to the case of the coefficient of discharge.

299

TABLE 5. EFFECT OF INITIAL CHAMBER VOLUME

U_0 (In³)	P_ℓ (KPSI)	P (KPSI)	V_B (FPS)	V_z (FPS)
0.6	24.3	19.0	2130	44.3
0.7	22.6	17.7	2100	41.7
0.8	22.0	17.2	2080	40.5
0.9	20.7	16.2	2050	38.7
1.0	1.6	15.6	1990	37.3

The effect of the area of the orifice is shown in Table 6. As area of the orifice is increased, more and more propellant is injected. Therefore, more hot gases are generated in the same period of time than before. Thus, the performance of the weapon is improved dramatically. It is better to promote combustion earlier in the ballistic cycle.

TABLE 6. EFFECT OF AREA OF THE ORIFICE

A_h (in²)	P_ℓ (KPSI)	P (KPSI)	V_B (FPS)	V_z (FPS)
0.010	11.9	9.52	1480	12.7
0.015	16.3	12.90	1780	23.2
0.020	20.3	16.00	2020	36.3
0.025	24.2	18.90	2150	51.6
0.030	28.9	22.50	2210	70.3

Table 7 shows the effect of the shot start pressure. In general, the shot start pressure is not a constant. It is a function of rifling design at the origin and engraving phenomena. As shot start pressure increases, so does the performance of the system because of rise in peak pressures. However, the variations are not severe as in other cases. To stabilize the combustion and also to improve the performance, it is better to have high shot start pressure.

TABLE 7. EFFECT OF SHOT START PRESSURE (C_d = 0.7)

$P_s(o)$ (PSI)	P_ℓ (KPSI)	P (KPSI)	V_B (FPS)	V_z (FPS)
2000	19.7	15.5	2000	37.6
4000	20.7	16.3	2040	40.1
5000	21.4	16.8	2060	40.4
6000	22.4	17.6	2080	40.9
7000	22.4	17.6	2080	40.9
8000	23.5	18.4	2130	42.2

VII. CONCLUSIONS

An accurate mathematical model is essential to characterize and predict the interior ballistics of a liquid propellant gun. A regenerative injection model is considered where the force derived from the combustion of a small portion of the propellant and a differential-area piston is utilized to inject the main charge into the chamber. The model involves a compressible liquid, injection fluid mechanics and the burning rate of propellant in addition to an imperfect gas law. The continuity, momentum, and energy equations were utilized in addition to two moving boundary conditions which were formulated by Newton's law. These were solved by numerical integration and digital computers. Model capability includes not only conventional interior ballistics results, but also the rate of propellant injection and the rate of propellant combustion as well as propellant distribution inside the chamber. The physical insight was obtained by neutron radiographic tests conducted at Oregon State University, image enhancement and data reduction at Los Alamos National Laboratory, and experimental test fixture firings by General Electric.

An excellent agreement is obtained between theory and experiment for the ratio of liquid pressure to gas pressure whenever the piston is in dynamic equilibrium during a portion of the ballistic cycle. The predicted results of piston motion and muzzle velocity are also in good agreement with experimental data. However, the agreement is not that good between theory and experiment for liquid and gas pressures. More experimentation and interaction between theory, model, and experimental data is desirable, especially in cases of propellant distribution and combustion. Long incubation periods have been observed either due to inadequate pyrotechnic charge or due to a characteristic of the liquid propellant system. The calculations indicate that droplets could reach the base of the projectile in the initial portion of the ballistic cycle. Droplet sizes of the order of 20 to 40 microns appear desirable in order to obtain reasonable ignition and combustion characteristics.

Limited parametric studies indicate that either lowering the discharge coefficient through the orifice, or increasing the chamber volume, yields lower peak pressures, lower muzzle velocities, and increase in time to reach peak pressure. If the area of the orifice is increased, lower muzzle velocity, higher piston velocity, an increase in peak pressure, and also a decrease in time to reach peak pressure will occur. If shot start pressure is increased, an increase both in muzzle velocity and peak pressure takes place. All findings indicate that this is the way to go in the future even if there are temporary setbacks. Therefore, the liquid propellant program should be continued without interruptions.
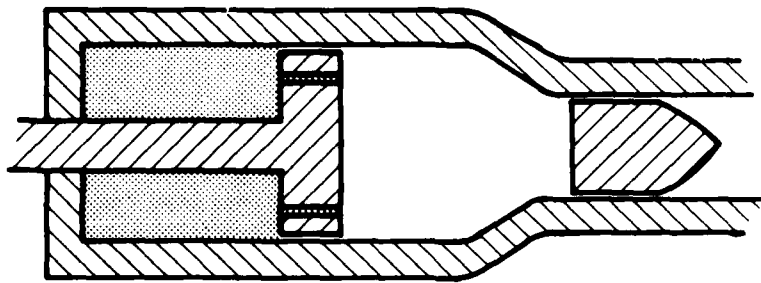
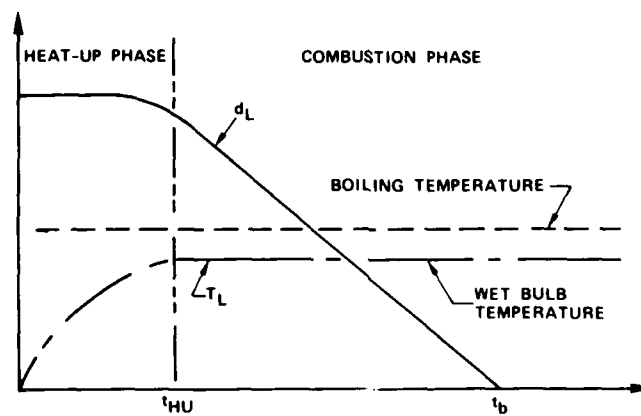FIGURE 1. SCHEMATIC OF REGENERATIVE LIQUID PROPELLANT SYSTEM
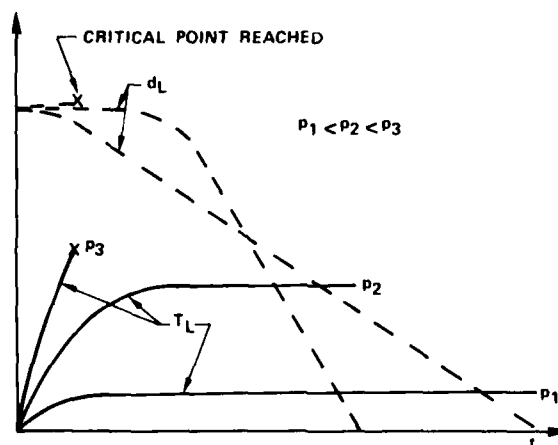

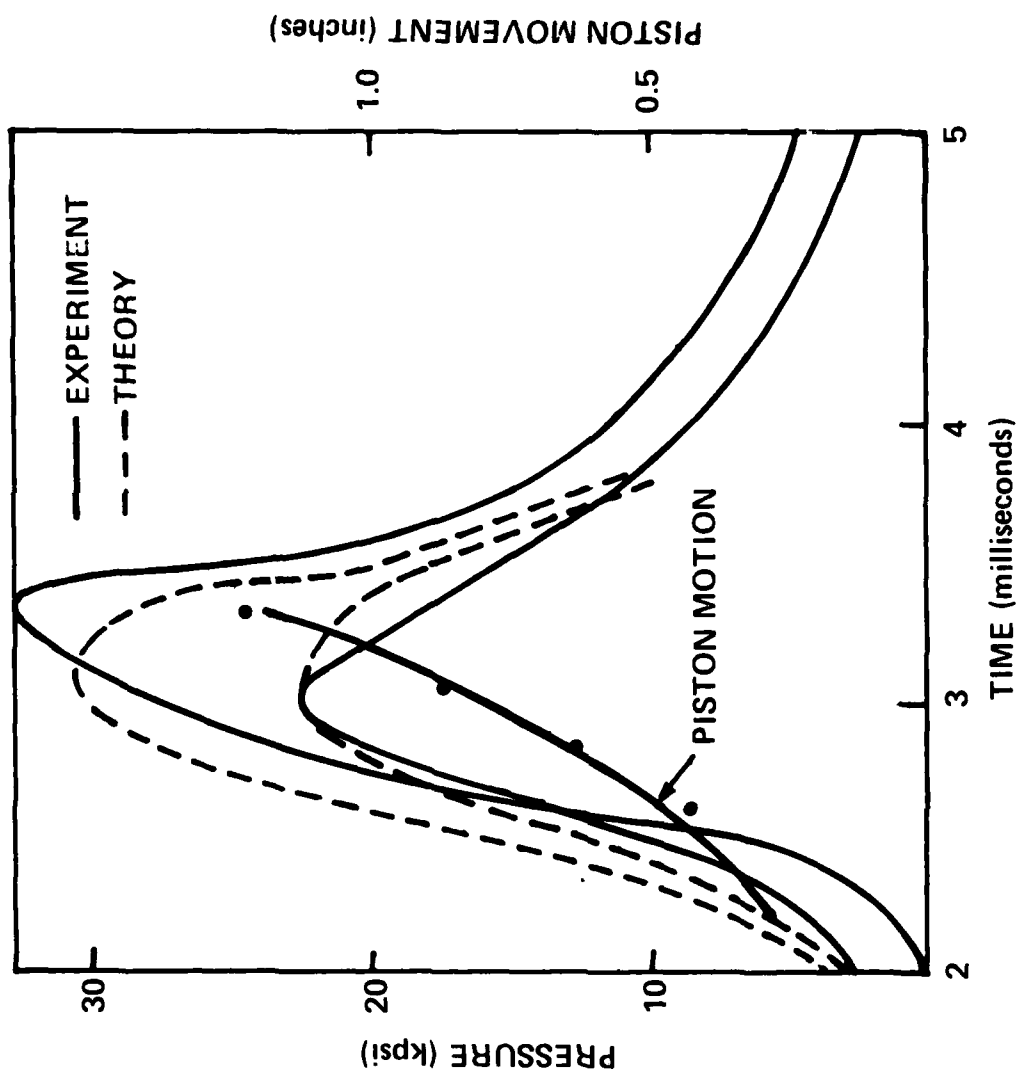FIGURE 2. PHYSICS OF A DROPLET

FIGURE 3. EFFECT OF PRESSURE ON DROPLET

FIGURE 4. COMPARISON OF THEORY AND EXPERIMENT

IMPLEMENTATION AND EVALUATION OF A MICROPROCESSOR
BASED SELF-TUNING CONTROL

PAK T. YIP
US ARMY ARMAMENT, MUNITIONS AND CHEMICAL COMMAND
US ARMY ARMAMENT RESEARCH AND DEVELOPMENT CENTER
FIRE CONTROL AND SMALL CALIBER WEAPON SYSTEMS LABORATORY
DOVER, NJ 07801

INTRODUCTION

The design of high performance weapon pointing system requires that
the weapon control system performs at or near its optimal level of perfor-
mance under unpredictable disturbance environments as well as situations
involving variations and degradation in system parameter values and com-
ponent characteristics. In the weapon pointing situation, gun pointing
commands are fed from the fire control computer to the digital controller
which has a high processing rate. The commands are treated as stepwise
constant inputs.

We use a deterministic digital control approach that involves a
stochastic formulation of the feedback control problem in which the gen-
eral state space equations are formulated into the state space innovation
form and then transformed into observable canonical form, which is then
readily converted to the required autoregressive moving average model.
The basic strategy is as follows:

1. Use the recursive Extended Least Squares (ELS) method [1, 2] to
identify the system paramters.

2. Construct an observer to provide the system state estimation.

3. Construct a feedback control law with pole asssignment [3, 4].

The design is simulated for program validation and then implemented
on an Intel 8086/87 microprocessor and tested with a laboratory inertia
wheel.

The weapon pointing system can be represented by the observable
canonical equation.

$$X(j+1) = AX(j) + Bu(j)$$

$$Y(j) = CX(j)$$

where

$$A = \begin{bmatrix} -a_1 & 1 & 0 & \cdots & 0 \\ \cdot & 0 & 1 & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & 1 \\ -a_n & 0 & \cdots & & 0 \end{bmatrix}$$

$$B = [b_1 \ldots b_n]^T$$

$$C = [1 \ 0 \ \ldots \ 0]$$

and $X(j)$ is the present state vector of the system, $Y(j)$ the measurement and n the number of system states.

Let $\hat{X}(j)$ be the estimate of $X(j)$. The innovation filter is

$$\hat{X}(j+1) = A\hat{X}(j) + Bu(j) + Ke(j)$$

$$\hat{Y}(j) = C\hat{X}(j)$$

$$e(j) = Y(j) - \hat{Y}(j)$$

where the gain K is to be properly chosen.

Using the shift operator z, the measurement equation is

$$Y(j) = C(I-Az^{-1})^{-1} Bz^{-1}u(j) + [1+C(I-Az^{-1})^{-1}Kz^{-1}] e(j)$$

After multiplying out the matrices, the measurement equation has an autoregressive moving average form.

306

$$a(z^{-1}) \, Y(j) = b(z^{-1}) \, u(j) + d(z^{-1}) \, e(j)$$

$$a(z^{-1}) = 1 + a_1 z^{-1} + \ldots + a_n z^{-n}$$

$$b(z^{-1}) = b_1 z^{-1} + \ldots + b_n z^{-n}$$

$$d(z^{-1}) = 1 + d_1 z^{-1} + \ldots + d_n z^{-n}$$

$$k_i = d_i - a_i$$

$$\forall i = 1, \ldots, n$$

where k_i is an element of the gain matrix K.

For a dead beat filter, $d_i = 0$ for all i.

Preparing for parameter identification, the measurement equation is arranged as

$$Y(j) = \Phi^T(j)\Theta + e(j)$$

$$\Theta = [a_1, \ldots, a_n, b_1, \ldots, b_n]^T$$

$$\Phi(j) = [-Y(j-1), \ldots, -Y(j-n), u(j-1), \ldots, u(j-n)]^T$$

Then the parameters can be estimated with the well established ELS method.

$$E(j+1) = y(j+1) - \Phi^T(j+1)\hat{\Theta}(j)$$

$$L(j+1) = \frac{P(j)\Phi(j+1)}{1 + \Phi^T(j+1)P(j)\Phi(j+1)}$$

$$P(j+1) = P(j) - \frac{P(j)\Phi(j+1)\Phi^T(j+1)P(j)}{1 + \Phi^T(j+1)P(j)\Phi(j+1)}$$

$$\hat{\Theta}(j+1) = \hat{\Theta}(j) + L(j+1)E(j+1)$$

307

Now the system observer can be formulated with the estimated A, B, and K matrices whose parameters are given by $\hat{\Theta}(j)$.

$$\hat{X}(j+1) = \hat{A} \, \hat{X}(j) + \hat{B}u(j) + \hat{K} \, [Y(j)-\hat{Y}(j)]$$

$$\hat{Y}(j) \quad = C\hat{X}(j)$$

Looking at the closed-loop system, the feedback control law is represented by

$$u(j) = r(j) - F_c^T \, T^{-1} \, \hat{X}(j)$$

where $r(j)$ is the reference input and the feedback vector $F_c = [f_{c1}, f_{c2}, \ldots f_{cn}]^T$

and

$$T^{-1} = T_1 T_2^{-1}$$

$$T_1 = [B \vdots AB \vdots \ldots \vdots A^{n-1}B]$$

$$T_2^{-1} = [C^T \vdots A^T C^T \vdots \ldots \vdots A^{T(n-1)}C^T]$$

The closed-loop response is

$$Y(j) = C(zI-A + BF_c^T \, T^{-1})^{-1} \, Br(j) + [1+C(zI-A+BF_c^T T^{-1})^{-1}K]e(j)$$

where

$$C(zI-A+BF_c^T T^{-1})^{-1}B = \frac{b_1 z^{-1} + \ldots + b_n z^n}{1+\alpha_1 z^{-1} + \ldots + \alpha_n z^{-n}}$$

$$\alpha_i = f_{ci}+a_i, \quad \forall i = 1, \ldots, n$$

$$a_{cL}(z^{-1}) = 1 + \alpha_1 z^{-1} + \ldots + \alpha_n z^{-n}$$

The characteristic equation

$$g_c(z) = \det [zI - A + BF_c^T T^{-1}]$$

$$= z^n + \sum_{i=1}^{n} (a_i + f_{ci}) z^{n-i}$$

$$= \prod_{i=1}^{n} (z - p_i)$$

The closed-loop poles p_i can be assigned as desired within the unit circle and then the f_{ci} of the feedback gain F_c can be readily computed.

Since the innovation process is zero mean Gaussian, the expectation $[e(j)] = 0$.

The expectation $[Y(j)] = \dfrac{b(1)}{a_{cL}(1)} r_c$

In order that the response tracks the input, we want expectation $[Y(j)] = r_c$.

Therefore, the feedback control law is modified as

$$u(j) = \frac{1 + \alpha_1 + \ldots + \alpha_n}{b_1 + \ldots + b_n} r(j) - F_c^T T^{-1} \hat{X}(j)$$

Combining the system identification, system observer and the pole assignment algorithm, we have a self-tuning control scheme for the weapon pointing system.

INERTIA WHEEL

The inertia wheel consists of two DC torque motors, which drive the inertia wheel and a resolver which measures the shaft angle. The inertia wheel is interfaced with electronics for check-out and computer command/control implementation. The resolver signal is amplified, demodulated, and converted to discrete format with an A/D converter. The discrete commands from a controller processor is converted to continuous signal with a D/A converter and the system is actuated with torque drive amplifiers.

If the friction is replaced by a constant, the open-loop response can be described by

$$y(s) = G[\frac{1}{S(S+\beta)}] \, u(s)$$

where G is the overall gain and β the open-loop pole.

Accounting for the effect of a zero order hold in the A/D converter, the open-loop response in z - domain is

$$y(z) = \frac{b_1 z^{-1} + b_2 z^{-2}}{1 + a_1 z^{-1} + a_2 z^{-2}} \, u(z)$$

Putting it into a state space form

$$y(j+1) = -a_1 y(j) + b_1 u(j) + [b_2 u(j-1) - a_2 y(j-1)]$$

Let

$$y_1(j) = y(j)$$

$$y_2(j) = b_2 u(j-1) - a_2(j-1)$$

Then

$$y_1(j+1) = -a_1 y_1(j) + b_1 u(j) + y_2(j)$$

$$y_2(j+1) = -a_2 y_1(j) + b_2 u(j)$$

Let
$$X(j) = \begin{bmatrix} y_1(j) \\ y_2(j) \end{bmatrix}$$

$$A = \begin{bmatrix} -a_1 & 1 \\ -a_2 & 0 \end{bmatrix}$$

310

YIP

$$B = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

$$C = [\ 1 \quad 0\ ]$$

Then $\quad X(j+1) = AX(j) + Bu(j)$

$$Y(j) = CX(j)$$

The a_1 and a_2 are system matrix parameters and the b_1 and b_2 are control weighting matrix parameters.

MICROPROCESSOR IMPLEMENTATION OF ALGORITHM

The microprocessor facility used in implementation of the self-tuning control algorithm is an Intel 8086/87 single board computer with 32K of random accessed memory and 8K of programmable read only memory. The 8086 runs with a 5-megahertz system clock. The 8087 coprocessor performs 32-bit floating point multiplication in 18 microseconds with 26 to 45 microseconds of overhead. For data conversion, an 8-channel 12-bit A/D and a 4-channel 12-bit D/A converters are used. The sampling time is programmable.
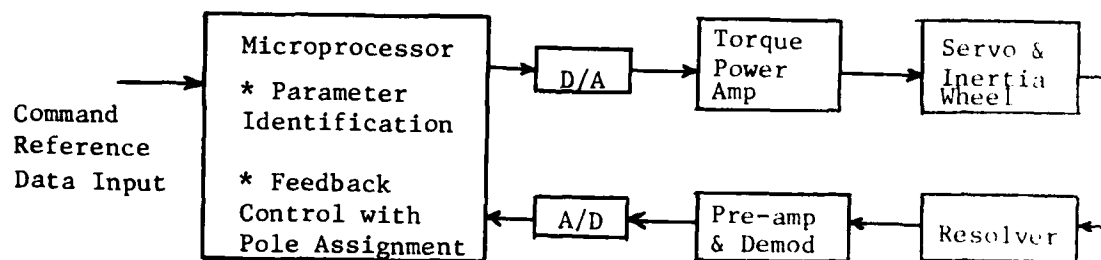
The system is set up as shown in figure 1.

Figure 1. Control System with a Digital Controller

With 20 samples per second processing rate and an irregular square wave input of approximately 0.8 Hz, the convergence of the system matrix

parameters a_i is very good with convergence time about 0.5 seconds as shown in figure 2. The control weighing matrix parameters b_i are functions of the resolver gain and torque amplifier gain which are linear when the input signals are small. The convergence time of the b-parameters is about 3 seconds with a bad initial value as shown in figure 3. The sensitivity of the identified parameters to their initial values with z-domain closed-loop poles at (0.4, 0.5) is shown in table 1.

When the assignments of closed-loop poles are changes, the identified parameter values are varied as shown in table 2. The bandwidth of the system is a function of the closed-loop pole assignments. It is interesting to see that the weight is increased on current control input when the pole assignments changed from (0.4, 0.5) to (0.2, 0.3).

The settling time of the closed-loop response for 20 samples per second processing rate is about 0.75 second with desired poles at (0.4, 0.5). The settling time for 50 samples per second processing rate is about 0.4 second with desired poles at (0.6, 0.7).
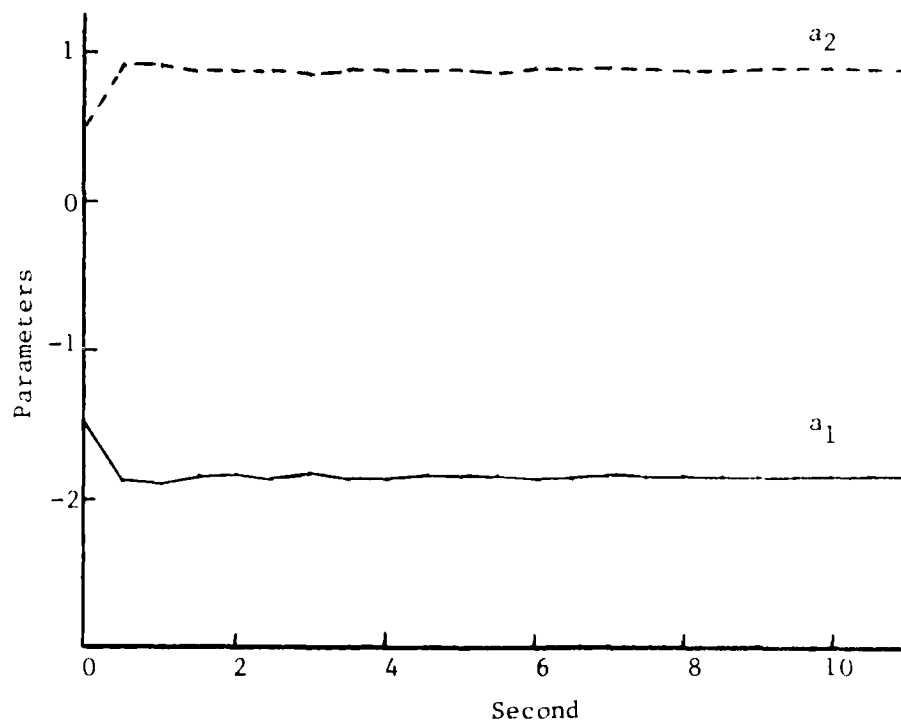

Figure 2. Identification of Parameters a_1 and a_2 with Square Wave Signal Input and with Closed Loop Poles at (0.4,0.5)
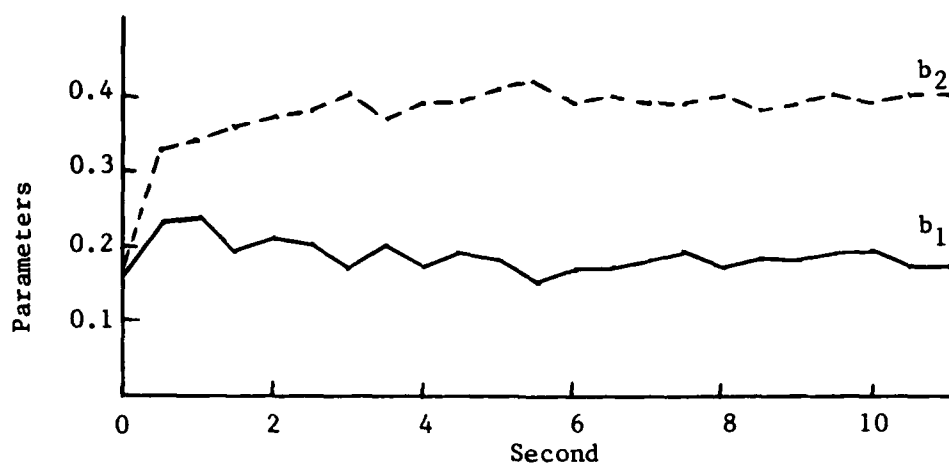
Figure 3. Identification of Parameters b_1 and b_2 with Square Wave Signal Input and with Closed Loop Poles at (0.4, 0.5)

Table 1. Sensitivity of Identified Parameters to its' Initial Value with Z-domain Closed-Loop Poles at (0.4, 0.5)

Initial Values				Identified Parameters			
a_1	a_2	b_1	b_2	a_1	a_2	b_1	b_2
-1.8	0.8	0.16	0.16	-1.85	0.86	0.172	0.416
-1.5	0.5	0.16	0.16	-1.85	0.86	0.174	0.400
-1.8	0.8	0.10	0.10	-1.85	0.86	0.162	0.396
-1.5	0.5	0.10	0.10	-1.85	0.86	0.177	0.406

Table 2. Sensitivity of Identified Parameters to Closed-Loop Pole Assignments

Closed-Loop Poles		Identified Parameters			
p_1	p_2	a_1	a_2	b_1	b_2
0.4	0.5	-1.85	0.86	0.172	0.416
0.2	0.3	-1.93	0.92	0.225	0.361

313

DISCUSSION

One of the purposes of designing a digital self-tuning controller is to process the gun command input signal which in general has a 10 per second sampling rate. The result of this study indicates that we need more than 100 per second processing rate in order to obtain the closed-loop response settling time less than 0.1 second. Fast multiplying chips and fast algorithm are required to accomplish this.

Because of the nonlinear elements in the physical system, the system parameters of a linear model are revealed as functions of closed loop pole assignments which define the bandwidth of the system.

In order to make full use of the digital self-tuning control algorithm, disturbance rejection should also be implemented in real time.

REFERENCES

1. L. Ljung, T. Soderstrom, and I. Gustarsson, "Counter Examples to General Convergence of a Commonly Used Recursive Identification Method," IEEE Trans. Automatic Control, AC-20, No. 5, pp. 643-652, 1975.

2. L. Ljung, "Convergence Analysis of Parameter Identification Methods," IEEE Trans. Automatic Control, AC-12, No. 5, pp. 770-783, 1978.

3. A.Y. Allidina and F.M. Hughers, "Generalized Self-Tuning Control with Pole Assignment," Proc. IEE, Vol. 127, Part D, No. 1, pp. 13-18, 1980.

4. Y.T. Tsay and L.S. Shieh, "State-Space Approach for Self-Tuning Feedback Control with Pole Assignment," Proc. IEE, Vol. 128, Part D, pp. 93-101, 1981.

END

FILMED

2-85

DTIC